



TUGAS AKHIR - ISI84853

**PENGEMBANGAN DASHBOARD ANALITIK
DESKRIPTIF UNTUK MEMANTAU PERTUMBUHAN
AKUN INSTAGRAM (*STUDI KASUS: PORTAL BERITA*)**

***DEVELOPMENT OF DESCRIPTIVE ANALYTICS
DASHBOARD TO MONITOR INSTAGRAM ACCOUNT
(CASE STUDY: NEWS PORTAL)***

AHMAD FAJAR PRASETYA
NRP 0521 12 4000 0118

Dosen Pembimbing
Rully Agus Hendrawan, S.Kom, M.Eng.

DEPARTEMEN SISTEM INFORMASI
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember
Surabaya 2019

TUGAS AKHIR - ISI84853

**PENGEMBANGAN DASHBOARD ANALITIK
DESKRIPTIF UNTUK MEMANTAU
PERTUMBUHAN AKUN INSTAGRAM (STUDI
KASUS: PORTAL BERITA)**

**AHMAD FAJAR PRASEYA
NRP 0521 12 4000 0118**

**Dosen Pembimbing
Rully Agus Hendrawan, S.Kom, M.Eng.**

**DEPARTEMEN SISTEM INFORMASI
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember
Surabaya 2019**

FINAL PROJECT - ISI 84853

**DEVELOPMENT OF DESCRIPTIVE
ANALYTICS DASHBOARD TO MONITOR
INSTAGRAM ACCOUNT (CASE STUDY:
NEWS PORTAL)**

**AHMAD FAJAR PRASETYA
NRP 0521 12 4000 0118**

**Supervisors
Rully Agus Hendrawan, S.Kom, M.Eng.**

**INFORMATION SYSTEMS DEPARTMENT
Information and Communication Technology Faculty
Tenth November Institut of Technology
Surabaya 2019**

LEMBAR PENGESAHAN

**PENGEMBANGAN DASHBOARD ANALITIK
DESKRIPTIF UNTUK MEMANTAU PERTUMBUHAN
AKUN INSTAGRAM (STUDI KASUS: PORTAL
BERITA)**

TUGAS AKHIR

Disusun Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada
Departemen Sistem Informasi
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember

Oleh:

AHMAD FAJAR PRASETYA

NRP. 0521 12 4000 0118

Surabaya, Januari 2019

**KETUA
DEPARTEMEN SISTEM INFORMASI**



Mahendrawathi ER, S.T., M.Sc., Ph.D

NIP.19660602 199203 1 002

LEMBAR PERSETUJUAN

**PENGEMBANGAN DASHBOARD ANALITIK
DESKRIPTIF UNTUK MEMANTAU PERTUMBUHAN
AKUN INSTAGRAM (STUDI KASUS: PORTAL
BERITA)**

TUGAS AKHIR

Disusun Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada
Departemen Sistem Informasi
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember

Oleh :

AHMAD FAJAR PRASETYA

NRP. 0521 12 4000 0118

Disetujui Tim Penguji : Tanggal Ujian: 16 Januari 2019
Periode Wisuda: 16-17 Maret 2019

Rully Agus Hendrawan, S.Kom, M.Eng (Pembimbing I)

Erma Suryani, S.T., M.T., Ph.D. (Penguji I)

Mahendrawathi ER, S.T., M.Sc., Ph.D (Penguji II)

**PENGEMBANGAN DASHBOARD ANALITIK
DESKRIPTIF UNTUK MEMANTAU PERTUMBUHAN
AKUN INSTAGRAM (STUDI KASUS: PORTAL
BERITA)**

Nama Mahasiswa : AHMAD FAJAR PRASETYA
NRP : 0521 12 4000 0118
Jurusan : DEPARTEMEN INFORMASI
FTIFK-ITS
Dosen Pembimbing 1 : Rully Agus Hendrawan S.Kom,
M.Eng

ABSTRAK

Instagram merupakan salah satu social network service (SNS) yang populer digunakan sebagai media beriklan beberapa waktu terakhir ini. Salah satu kategori linimasa populer pada intagram adalah kategori berita. Beberapa akun portal berita yang ada pada instagram menarik untuk menjadi studi kasus. Detikcom menjadi salah satu akun portal berita dengan followers tertinggi dengan nilai followers mencapai lebih dari satu juta akun yang mengikuti berita terbaru dari detikcom. Sementara kompascom dan tribunnews mempunyai tingkat followers yang lebih rendah dari detikcom dengan kisaran angka enam ratus ribu akun yang mengikuti kedua akun portal berita pada instagram tersebut. Sindonews sendiri mempunyai followers paling rendah dengan hanya ada sekira enam puluh ribu akun yang mengikuti akun portal berita ini.

Pada penelitian kali ini digunakan DSDM sebagai metodologi pengembangan dashboard. Proses pembuatan dashboard sendiri dibagi menjadi scraping menggunakan

python yang datanya disimpan di basis data MySQL yang kemudian ditampilkan pada Dashboard yang dibuat dengan tableau 9.2. Sebelum menyimpan data pada basis data sendiri hasil dari scraping tersebut berupa data yang tersimpan pada ekstensi json. Data json ini kemudian dimasukan kedalam basis data MySQL menggunakan script PHP sederhana.

Hasil yang diharapkan dari penelitian ini adalah dashboard yang mampu menampilkan pertumbuhan jumlah followers dan tingkat keaktifan dari sebuah akun portal berita. Tingkat keaktifan ini dinilai dari jumlah posts yang dibuat oleh akun portal berita dan jumlah likes pada post yang dibuat akun portal berita.

Kata kunci: SNS, influencer, rating, post, hashtag, Instagram, dashboard

**DEVELOPMENT OF DESCRIPTIVE ANALYTICS
DASHBOARD TO MONITOR INSTAGRAM ACCOUNT
(CASE STUDY: NEWS PORTAL)**

Name : AHMAD FAJAR PRASETYA
NRP : 0521 12 4000 0118
Departement : INFORMATION SYSTEM FTIFK-ITS
Supervisor : Rully Agus Hendrawan S.Kom, M.Eng

ABSTRACT

Instagram is one of the social network services (SNS) that is popularly used as a medium for advertising in recent times. One popular timeline category on intagram is the news category. Some news portal accounts available on Instagram are interesting to be a case study. Detikcom is one of the news portal accounts with the highest followers with followers reaching more than one million accounts following the latest news from AFP. While Kompascom and Tribunnews have lower followers than AFP with a range of six hundred thousand accounts that follow both news portal accounts on the Instagram account. Sindonews itself has the lowest followers with only around sixty thousand accounts that follow this news portal account.

In this research, DSDM is used as a dashboard development methodology. The process of creating the dashboard itself is divided into scraping using python whose data is stored in the MySQL database which is then displayed on the Dashboard created with tableau 9.2. Before storing data in its own database the results of the scraping are in the form of data stored in the json extension. This json data is then entered into the MySQL database using a simple PHP script.

The expected results of this research are dashboards that are able to display the growth in the number of followers and the

level of activity of a news portal account. This level of activity is judged by the number of posts made by the news portal account and the number of likes in the post made by the news portal account.

Keywords: SNS, influencers, rating, post, hashtag, Instagram, dashboard

KATA PENGANTAR

Alhamdulillah atas karunia, rahmat, barakah, dan jalan yang telah diberikan Allah SWT selama ini sehingga penulis mendapatkan kelancaran dalam menyelesaikan tugas akhir dengan judul:

PENGEMBANGAN DASHBOARD ANALITIK DESKRIPTIF UNTUK MEMANTAU PERTUMBUHAN AKUN INSTAGRAM (*STUDI KASUS: PORTAL BERITA*)

Terima kasih atas pihak-pihak yang telah mendukung, memberikan saran, motivasi, semangat, dan bantuan baik materi maupun spiritual demi tercapainya tujuan pembuatan tugas akhir ini. Secara khusus penulis akan menyampaikan ucapan terima kasih yang sedalam-dalamnya kepada:

1. Allah SWT yang senantiasa melimpahkan berkat dan rahmat-Nya selama penulis melaksanakan Tugas Akhir.
2. Bapak Rully Agus Hendrawan, S.Kom., M.Sc selaku dosen pembimbing yang meluangkan waktu, memberikan ilmu, petunjuk, dan motivasi untuk kelancaran Tugas Akhir ini.
3. Ibu Erma Suryani, S.T., M.T., Ph.D. dan Ibu Mahendrawathi ER, S.T., M.Sc., Ph.D selaku dosen penguji yang telah memberikan masukan untuk perbaikan tugas akhir ini.
4. Orang tua penulis dan saudara kandung penulis telah mendokan dan mendukung.
5. Seluruh dosen Departemen Sistem Informasi ITS yang telah memberikan ilmu yang sangat berharga bagi penulis.

6. Rekan-rekan SOLA12IS yang telah berjuang bersama dalam menjalani perkuliahan di Departemen Sistem Informasi ITS.
7. Berbagai pihak yang membantu dalam penyusunan Tugas Akhir ini dan belum dapat disebutkan satu per satu dengan dukungan, semangat dan kebersamaan.

Penyusunan laporan ini masih jauh dari sempurna, untuk itu saya menerima adanya kritik dan saran yang membangun untuk perbaikan di masa mendatang. Semoga buku tugas akhir ini dapat memberikan manfaat pembaca.

Surabaya, Januari 2019
Penulis,

(Ahmad Fajar Prasetya)

DAFTAR ISI

LEMBAR PENGESAHAN....	Error! Bookmark not defined.
LEMBAR PERSETUJUAN...	Error! Bookmark not defined.
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah	3
1.3. Batasan Pengerjaan Tugas Akhir	3
1.4. Tujuan Tugas Akhir	3
1.5. Manfaat Tugas Akhir	3
1.6. Relevansi.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1. Studi Sebelumnya	5
2.2. Dasar Teori	8
2.2.1. Instagram.....	9
2.2.2. Metadata.....	9
2.2.3. Database Schema	9
2.2.4. Python	10
2.2.5. <i>Web Scraping</i>	11
2.2.6. MySQL	11
2.2.7. Rapid Application Development (RAD).....	11
2.2.8. Visualisasi Dashboard.....	12
2.2.9. Tableau 9.2.....	18
2.2.10. Json	18
BAB III METODOLOGI PENELITIAN.....	21
3.1. Metodologi Penelitian.....	21
3.2. Tahap Pelaksanaan.....	21
3.2.1. Studi Literatur	21
3.2.2. Kajian Bisnis.....	22
3.2.3. Pembuatan Model Fungsional.....	22
3.2.4. Desain Sistem dan Pembuatan Basis Data.....	22

3.2.5.	Analisis dan Visualisasi Data	23
3.2.6.	Penyusunan Laporan Tugas Akhir	23
3.3.	Rangkuman Metodologi.....	23
BAB IV	PENGEMBANGAN APLIKASI SCRAPING	25
4.1.	Studi Kasus	25
4.1.1.	Observasi halaman <i>web</i> instagram	25
4.2.	Pembuatan <i>Script Scraping</i>	31
4.2.1.	<i>Library</i> Python instagram-profilecrawl.....	31
4.2.2.	<i>Flowchart Script Library</i>	33
4.2.3.	Pengumpulan Data	38
4.3.	Desain Basis Data	39
4.3.1.	Pembuatan Skema Basis Data	40
4.3.2.	Pembuatan Basis Data	45
4.3.3.	Pengisian Basis Data	49
BAB V	UJI COBA	53
5.1.	Uji Coba	53
5.2.	Spesifikasi	54
5.2.1.	Perangkat Keras.....	55
5.2.2.	Perangkat Lunak.....	55
BAB VI	VISUALISASI DAN ANALISIS DATA	57
6.1.	Pengolahan Data Pada <i>Tableau 9.2</i>	57
6.1.1.	Memasukan Data Pada <i>Tableau 9.2</i>	57
6.1.2.	Visualisasi Data.....	59
6.1.3.	Grafik Perbandingan <i>Followers</i> dan <i>Post</i>	64
6.2.	<i>Followers Growth Analytics</i>	69
6.3.	Analisis Pertumbuhan <i>Followers</i> dan <i>Post</i>	70
BAB VII	KESIMPULAN DAN SARAN	79
7.1.	Kesimpulan	79
7.2.	Saran	80
DAFTAR PUSTAKA		81
LAMPIRAN A SOURCE CODE		1

DAFTAR GAMBAR

Gambar 2.1 Contoh <i>bar chart</i>	12
Gambar 2.2 Contoh <i>line chart</i>	13
Gambar 2.3 Contoh <i>pie chart</i>	14
Gambar 2.4 contoh <i>nap chart</i>	14
Gambar 2.5 Contoh <i>heat map</i>	15
Gambar 2.6 Contoh <i>highlight table</i>	16
Gambar 2.7 Contoh <i>bullet chart</i>	16
Gambar 2.8 Contoh histogram	17
Gambar 2.9 Contoh <i>scatter plots</i>	18
Gambar 3.1 Metodologi penelitian.....	21
Gambar 4.1 <i>Script</i> json halaman utama.....	26
Gambar 4.2 Sumber data halaman <i>followers</i>	28
Gambar 4.3 Sumber data halaman <i>post</i>	30
Gambar 4.4 Informasi <i>library</i> yang dipakai.....	32
Gambar 4.5 <i>Flowchart</i> crawl_profile.py.....	34
Gambar 4.6 <i>Flowchart</i> cli_helper.py	35
Gambar 4.7 <i>Flowchart</i> datasavers.py	35
Gambar 4.8 <i>Flowchart</i> extractor.py.....	36
Gambar 4.9 Skema tabel <i>user_account</i>	40
Gambar 4.10 Skema tabel <i>posts</i>	42
Gambar 4.11 Skema tabel <i>comments</i>	43
Gambar 4.12 Setiap tabel pada basis data	45
Gambar 4.13 Pembuatan basis data <i>crawl</i>	46
Gambar 4.14 <i>Script</i> pembuatan tabel <i>user_account</i>	46
Gambar 4.15 <i>Script</i> pembuatan tabel <i>posts</i>	46
Gambar 4.16 Struktur tabel <i>posts</i>	47
Gambar 4.17 Menambah <i>foreign key</i> <i>username</i>	47
Gambar 4.18 Struktur tabel terbaru	47
Gambar 4.19 <i>Script</i> pembuatan tabel <i>comments</i>	48
Gambar 4.20 Struktur tabel awal <i>comments</i>	48
Gambar 4.21 <i>Script</i> penambahan <i>fk</i> <i>username</i>	48
Gambar 4.22 <i>Script</i> penambahan <i>fk</i> url.....	49
Gambar 4.23 Struktur akhir tabel <i>comments</i>	49

Gambar 4.24 Basis data <i>crawl</i>	49
Gambar 4.25 <i>Flowchart</i> memasukan data	50
Gambar 4.26 Tampilan memasukan data json ke basis data ..	51
Gambar 6.1 Koneksi ke server basis data.....	57
Gambar 6.2 Potongan tiap tabel pada tableau	58
Gambar 6.3 Data yang dianalisis pada tableau	58
Gambar 6.4 Perbandingan jumlah <i>followers</i> dan <i>posts</i>	59
Gambar 6.5 Perbandingan jumlah <i>followers</i> dan <i>posts</i> detikcom	60
Gambar 6.6 Perbandingan jumlah <i>followers</i> dan <i>posts</i> kompascom.....	61
Gambar 6.7 Perbandingan jumlah <i>followers</i> dan <i>posts</i> sindonews	62
Gambar 6.8 Perbandingan jumlah <i>followers</i> dan <i>posts</i> tribunnews	63
Gambar 6.9 Perbandingan jumlah <i>followers</i>	64
Gambar 6.10 Perbandingan jumlah <i>post</i>	65
Gambar 6.11 Perbandingan perubahan <i>post</i> dan <i>followers</i>	66
Gambar 6.12 Perbandinan jumlah <i>likes</i> tiap akun	68
Gambar 6.13 Dashboard <i>followers</i> growth analytics.....	70
Gambar 6.14 Jumlah <i>post</i> yang dibuat	71
Gambar 6.15 Jumlah <i>followers</i> yang didapat	72
Gambar 6.16 Prosentase pertumbuhan <i>followers</i>	73
Gambar 6.17 Prosentase perubahan <i>followers</i> dan jumlah perubahan <i>followers</i>	73
Gambar 6.18 Perbandingan jumlah <i>followers</i> dan <i>post</i>	74
Gambar 6.19 Pertumbuhan akun dengan basis <i>folllowers</i>	76

DAFTAR TABEL

Table 2.1 Penelitian sebelumnya.....	5
Table 3.1 Input dan Output Metodologi.....	24
Table 4.1 Tipe data halaman utama <i>script</i> json.....	27
Table 4.2 Tipe data halaman utama <i>script</i> html	28
Table 4.3 Data pada halaman <i>post</i>	30
Table 4.4 <i>Dependencies</i> sistem <i>crawling</i>	33
Table 4.5 Pengambilan elemen halaman html instagram.....	36
Table 4.6 Struktur Data Json	38
Table 4.7 Perancangan tabel <i>user_account</i>	41
Table 4.8 Perancangan tabel <i>posts</i>	42
Table 4.9 Perancangan tabel <i>comments</i>	44
Table 5.1 Uji Coba	53

BAB I

PENDAHULUAN

Dalam bab pendahuluan ini akan menjelaskan mengenai latar belakang masalah, perumusan masalah, batasan masalah, tujuan tugas akhir, dan manfaat dari kegiatan tugas akhir. Berdasarkan uraian pada bab ini diharapkan mampu memberi gambaran umum permasalahan dan pemecahan masalah pada tugas akhir.

1.1. Latar Belakang

Dengan maraknya penggunaan gawai di masyarakat, hal ini menyebabkan pertumbuhan aplikasi *Social Network Service* (untuk selanjutnya disebut SNS). SNS dapat didefinisikan halaman *web* individual yang memungkinkan pengguna untuk menjalin hubungan sosial dengan mengumpulkan informasi dan membaginya dengan orang lain, baik spesifik maupun tidak spesifik secara *online* [1]. Beberapa SNS seperti *facebook*, *Instagram*, *Twitter* dan lain sebagainya memungkinkan penggunaannya untuk mengunggah foto, video maupun teks. Hal ini memungkinkan terjadinya interaksi yang lebih intim antar pengguna SNS.

A picture tell a thousand words. Salah satu aplikasi SNS yang populer belakangan ini adalah *Instagram*, menurut data dari Statista sampai September 2017 terdapat 800 juta pengguna *Instagram* tiap bulannya [2]. Terdapat hal menarik dari *instagram* yang membuat peneliti memutuskan untuk membuat topik penelitian ini jika dibandingkan dengan SNS lain selain karena jumlah penggunanya yang besar. Menurut Omnicoreagency [3], video yang diunggah di *instagram* mempunyai tingkat *engagement* yang lebih tinggi jika dibandingkan dengan SBS lain. Selain itu *Intsgram* merupakan *mobile SNS* memungkinkan penggunaannya untuk berbagi foto maupun video disertai teks atau *caption* hal ini menyebabkan mudahnya dilakukan pemasaran karena ketika pengguna membuka *instagram* yang mereka cari adalah foto dan video. Salah satu fitur fungsional dari *Instagram* adalah *timeline*, fitur ini memungkinkan pengguna untuk mendapat *update* terbaru

dari akun yang telah pengguna ikuti. Fitur lainnya adalah *search*, fitur ini dibuat khusus untuk menunjukkan unggahan yang menarik untuk pengguna berdasar dari sejarah akun pengguna. Selain itu terdapat pula fitur *story*, yang menampilkan aktivitas akun yang pengguna ikuti. Hal ini menyebabkan *Instagram* salah satu *platform* yang baik untuk melakukan promosi produk. Dalam sebuah *post* di *instagram* ada yang menggunakan *hashtag* untuk memberikan keterangan pada foto atau video yang mereka unggah. Juga menurut omnicoagency, *post* yang diberi keterangan dengan *hashtag* walaupun cuma satu *hashtag* memberikan tingkat interaksi 12,6 persen lebih tinggi.

Untuk penyedia layanan dan produk, membuat akun pada *Instagram* dengan pengikut yang banyak adalah impian. Tapi bagaimana akun tersebut bila dibandingkan dengan akun pesaing. Pengikut yang banyak adalah modal utama sebuah akun. Ada sebuah fenomena di *Instagram* untuk korporasi, mereka menggunakan akun orang lain seperti *public figures* untuk mempromosikan produk atau layanannya. Mereka disebut *selebgram* atau *influencer* bagi korporasi tersebut. Memilih *influencer* dan membuat mereka *engaged* ke demografi yang ingin dicapai merupakan syarat sukses *marketing* pada *Instagram*.

Portal berita merupakan salah satu kategori dalam lini masa terpopuler pada aplikasi *instagram* [4]. Hal ini menjadikan akun portal berita sebagai salah satu kategori yang potensial untuk pemilik bisnis membuat iklan pada akun tersebut.

Untuk mengetahui kelayakan sebuah akun portal berita untuk digunakan sebagai media promosi produk dapat dilihat dari beberapa kriteria antara lain pertumbuhan jumlah *followers*, jumlah *posts* dan jumlah *likes post* yang dibuat oleh akun portal berita. Jumlah *followers* dapat bertambah karena keaktifan dari akun tersebut dalam membuat *post* baru. Dengan mempunyai jumlah *followers* lebih banyak *post* yang dibuat akun tersebut

akan tersebar ke lebih banyak orang karena algoritma pada instagram memungkinkan hal ini untuk terjadi.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan pada bagian sebelumnya, maka rumusan masalah yang akan diselesaikan pada Tugas Akhir ini adalah sebagai berikut:

1. Bagaimana pertumbuhan sebuah akun instagram portal berita?
2. Bagaimana perubahan jumlah *followers* akun instagram?
3. Bagaimana perubahan jumlah *posts* dan *likes* akun instagram?
4. Bagaimana informasi pertumbuhan akun instagram dalam bentuk grafik?

1.3. Batasan Pengerjaan Tugas Akhir

Sesuai dengan deskripsi permasalahan yang telah dijelaskan diatas, adapun batasan permasalahan dari penyelesaian tugas akhir ini adalah sebagai berikut:

1. *Scraping* dilakukan dengan *python 3*.
2. *Dashboard* dibuat dengan *tableau 9.2*

1.4. Tujuan Tugas Akhir

Berdasarkan hasil perumusan masalah dan batasan masalah yang telah disebutkan sebelumnya, maka tujuan untuk membuat *dashboard* untuk mengetahui bagaimana pertumbuhan sebuah portal berita pada instagram berdasar jumlah *post*, *likes* dan *followers*.

1.5. Manfaat Tugas Akhir

Manfaat dari tugas akhir ini adalah:

1. Bagi akademisi, mampu membuat *dashboard* yang menampilkan pertumbuhan jumlah *followers* dan perbandingan jumlah *post*.
2. Bagi pemilik bisnis, mampu menentukan portal berita yang menguntungkan untuk dipasang iklan.

1.6. Relevansi

Tugas akhir ini berkaitan dengan mata kuliah Interaksi Manusia Komputer, Algoritma dan Pemrograman, Pemrograman Berorientasi Objek, Pemrograman Berbasis Web, Desain Basis Data, Analisis dan Desain Perangkat Lunak, Pemrograman Integratif dan Konstruksi Pengembangan Perangkat Lunak.

BAB II TINJAUAN PUSTAKA

Bagian ini akan memberikan penjelasan mengenai penelitian maupun studi literatur sebelumnya yang berkaitan dan dijadikan sebagai acuan selama pengerjaan tugas akhir, serta landasan teori yang berkaitan dengan tugas akhir yang dapat membantu pemahaman selama pengerjaan tugas akhir ini.

2.1. Studi Sebelumnya

Selama pengerjaan tugas akhir ini, terdapat beberapa penelitian yang telah dilakukan sebelumnya yang dapat dijadikan sebagai bahan kajian maupun referensi untuk studi literatur. Penelitian tersebut lalu dikaji untuk dilihat dari gambaran umum, tujuan hasil, dan keterkaitannya dengan penelitian tugas akhir ini. Hasil dari kajian tersebut dapat kita lihat pada tabel berikut ini.

Table 2.1 Penelitian sebelumnya

Penelitian	1
Judul Penelitian	<i>Multivariate Features Based Instagram Post Analysis to Enrich User Experience</i> [5]
Metode	<i>Non-negative matrix factorization dan latent Dirichlet allocation algorithm</i>
Penulis	Vatsala Mittal, Aastha Kaul, Santoshi Sen Gupta, Anuja Arora,
Tanggal Penerbitan	2017
Keterkaitan Tugas Akhir	Penelitian ini berusaha menganalisis interaksi dan <i>engagement</i> antara pengguna <i>instagram</i> . Untuk memahami tingkat <i>engagement</i> pengguna <i>instagram</i> terhadap sebuah <i>post</i> di <i>instagram</i> , perlu dilihat bagaimana properti interaksi

	<p>pengguna yang menonjol dan properti foto. Properti interaksi pengguna antara lain <i>hashtag</i>, waktu <i>post</i> foto, sementara itu properti foto adalah fitur foto atau konteks seperti <i>filter</i> yang ada pada foto tersebut. Terdapat empat analisis yang digunakan untuk melakukan komputasi, keempatnya adalah <i>time based analysis (TBA)</i>, <i>image filter analysis (IFA)</i>, <i>image hashtags analysis (IHA)</i> and <i>image categorization analysis (ICA)</i>.</p>
Penelitian	2
Judul Penelitian	<i>Self-Branding on Social Media: An Analysis of Style Bloggers on Instagram</i> [6]
Metode	<i>Content analysis method</i>
Penulis	Rendan Liu, Ayoung Suh
Tanggal Penerbitan	2017
Keterkaitan Tugas Akhir	<p>Melakukan analisis terhadap 243 sampel <i>post</i> dari 10 profil <i>blogger</i> gaya <i>instagram</i>. Hasil dari analisis tersebut mengindikasikan walaupun sebgayaan <i>blogger</i> gaya masih mengikuti kecantikan tradisional dan kriteria hierarkis. Sosial media memberikan kebebasan bagi mereka untuk membuat <i>brand</i> diri mereka dan memasarkan diri mereka dengan semakin meningkatnya kesadaran akan hak hak wanita. Penelitian ini juga menemukan bahwa kebanyakan pengikut mereka terpengaruh dan</p>

	menjadi terikat secara emosional kepada <i>blogger</i> tersebut.
Penelitian	3
Judul Penelitian	<i>Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users</i> [7]
Metode	Metode riset kualitatif: wawancara mendalam
Penulis	Elmira Djafarova, Chloe Rushworth
Tanggal Penerbitan	2017
Keterkaitan Tugas Akhir	Penelitian ini menemukan bahwa selebritas di <i>instagram</i> masih memberi pengaruh pada perilaku pembelian pada wanita muda pengguna <i>instagram</i> . Meski demikian selebritas non tradisional seperti <i>bloggers</i> , <i>youtber</i> dan selebgram memiliki pengaruh yang lebih besar karena partisipan menganggap mereka lebih kredibel dan lebih terkait dengan partisipan. Wanita pengguna <i>instagram</i> lebih tanggap dan lebih memilih untuk mengikuti profil <i>instagram</i> yang menggambarkan positif dan ulasan yang mendorong.
Penelitian	4
Judul Penelitian	<i>What we instagram : a first analysis of instagram photo content and user types</i> [8]

Metode	<i>Data crawling</i> dan melakukan kategorisasi foto <i>instagram</i> secara manual
Penulis	Yuheng Hu, Lydia Manikonda, Subbarao Kmabhampati
Tanggal Penerbitan	2014
Keterkaitan Tugas Akhir	Merupakan salah satu riset paling awal mengenai Instagram. Berisi hasil <i>crawling</i> melalui API yang dimiliki <i>Instagram</i> . <i>Paper</i> ini berusaha melakukan analisis terhadap isi foto dan tipe pengguna <i>instagram</i> . Dalam <i>paper</i> ini dilakukan kategorisasi terhadap jenis foto yang diunggah oleh pengguna <i>instagram</i> . Kategorisasi ini dilakukan secara manual, dan didapat delapan kategori untuk jenis foto yang diunggah, kategori tersebut adalah teman, makanan, perkakas, foto dengan keterangan, hewan peliharaan, aktifitas, <i>selfie</i> dan <i>fashion</i> . Berdasar konten yang diunggah oleh pengguna, pada analisis ini ditemukan terdapat lima tipe pengguna <i>instagram</i> . Dalam penelitian ini juga tidak ditemukan adanya hubungan langsung antara jumlah <i>followers</i> dan tipe pengguna berdasarkan foto yang dibagi.

2.2. Dasar Teori

Pada bab ini dijelaskan mengenai teori – teori yang digunakan dalam pengerjaan tugas akhir ini.

2.2.1. Instagram

Instagram diluncurkan pada Oktober 2010 sebagai sarana berbagi foto dan SNS untuk sistem operasi iOS. Versi untuk android baru diluncurkan dua tahun kemudian pada tahun 2012, yang diikuti diluncurkannya versi website untuk *instagram* pada tahun yang sama. Sekarang *Instagram* dimiliki oleh facebook.

Instagram memberikan sarana bagi pengguna untuk mengunggah foto atau video pada layanan yang mereka berikan. Pengguna juga dapat memberi *filter* dan lokasi untuk foto yang mereka unggah. *Instagram* dapat menghubungkan topik yang sama dengan memasang *hashtag* pada *post* yang telah diunggah. Cara untuk berkomunikasi lewat *Instagram* adalah dengan "*likes*" dan meninggalkan komentar pada *post* pengguna lain [9].

2.2.2. Metadata

Metadata secara umum didefinisikan sebagai data terstruktur dari sebuah data dan informasi [9]. Dalam sudut pandang basis data, Metadata adalah basis data yang berisi informasi tentang struktur data, arti data, penggunaan data, aturan kualitas data, dan informasi lain tentang data [11].

Metadata dibutuhkan untuk menggambarkan dan menjelaskan data dan gudang data / data warehouse dalam artian struktur data dan proses bagi pengguna. Definisi Metadata dilakukan untuk menghindari kesalah pahaman akan makna dari suatu kolom tertentu dimana pemetaan metadata juga membantu pengguna untuk memahami dampak dari perubahan atau penambahan kolom baru pada basis data.

2.2.3. Database Schema

Database Schema merupakan sebuah struktur kerangka yang mewakili pandangan logis dari basis data. *Database Schema* mendefinisikan bagaimana data diatur dan bagaimana keterkaitan hubungan di antara mereka . *Database Schema*

memberikan deskripsi hubungan logis secara lengkap dari basis data yang meliputi rincian nama-nama dari tiap *field* (atau atribut atau kolom) dan tipe dari tiap *field*.

Pada DBMS, Database Schema menjelaskan struktur formal language yang ada pada DBMS dan dapat dikatakan sebagai penjelasan dari database itu sendiri. Database Schema menunjukkan bagaimana entitas yang ada dalam database berhubungan satu sama lain, termasuk tables, views, stored procedures, dan banyak lagi. Database Schema tidak boleh sering berubah karena akan berpengaruh terhadap data yang disimpan di dalam database [12].

2.2.4. Python

Python adalah sebuah bahasa pemrograman interpretatif multiguna dengan filosofi perancangan yang berfokus pada tingkat keterbacaan kode. Python diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif [13]. Bahasa pemrograman ini diciptakan oleh Guido van Rossum pada tahun 1990 silam. Python mendukung multi paradigma pemrograman, utamanya namun tidak dibatasi pada pemrograman berorientasi objek, pemrograman imperatif, dan pemrograman fungsional. Salah satu fitur yang tersedia pada python adalah sebagai bahasa pemrograman dinamis yang dilengkapi dengan manajemen memori otomatis. Seperti halnya pada bahasa pemrograman dinamis lainnya, python umumnya digunakan sebagai bahasa skrip meski pada praktiknya penggunaan bahasa ini lebih luas mencakup konteks pemanfaatan yang umumnya tidak dilakukan dengan menggunakan bahasa skrip. Python dapat digunakan untuk berbagai keperluan pengembangan perangkat lunak dan dapat berjalan di berbagai platform sistem operasi.

2.2.5. *Web Scraping*

Web scraping merupakan metode untuk mendapatkan informasi dari *web* dengan menuliskan program computer untuk melakukan apa yang kita lakukan secara manual seperti memilih, menyalin dan melakukan *paste* informasi pada *web* menjadi otomatis [14].

2.2.6. MySQL

MySQL merupakan salah satu database besar yang menawarkan skalabilitas, keluwesan dan kinerja yang tinggi. Database ini didukung oleh banyak bahasa pemrograman dan meluas di berbagai platform. MySQL merupakan *database* paling populer di dunia [15]. Dengan performa, reliabilitas dan kemudahan penggunaan, MySQL menjadi pilihan untuk aplikasi berbasis web.

2.2.7. Rapid Application Development (RAD)

James Martin, orang yang pertama menggunakan istilah ini menuliskan dalam bukunya [16], "Rapid Application Development (RAD) adalah siklus pengembangan yang dirancang untuk memberikan pengembangan yang lebih cepat dan hasil berkualitas lebih tinggi daripada yang dicapai dengan siklus hidup tradisional. Hal ini dirancang untuk mengambil keuntungan maksimal dari perangkat lunak pengembangan yang kuat yang telah berkembang baru baru ini."

Salah satu metode pengembangan aplikasi dengan RAD adalah DSDM (*Dynamic System Development Method*) yang dimiliki konsorsium DSDM. Salah satu keunggulan DSDM adalah pengujian terintegrasi selama iterasi pengembangan aplikasi. Selain itu semua perubahan dalam pengembangan dapat dibalikkan, yang artinya kemampuan untuk kembali ke versi sebelumnya dari aplikasi merupakan fitur penting dari DSDM [17].

Terdapat lima fase pengerjaan dalam DSDM, mereka antara lain:

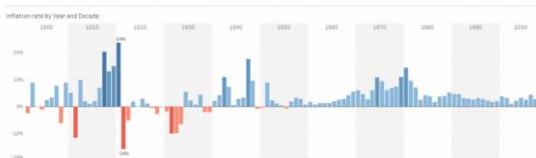
1. Studi kelayakan: Fase ini mempertimbangkan kelayakan proyek dalam hal bisnis dan teknis serta kesesuaian proyek dengan pendekatan RAD.
2. Studi bisnis: Fase ini mendefinisikan fungsionalitas tingkat tinggi dan entitas bisnis utama yang terpengaruh.
3. Model iterasi fungsional: Fase ini digunakan untuk membangun dan mendemonstrasikan fungsi yang dibutuhkan menggunakan prototipe kerja.
4. Perancangan sistem dan pembuatan iterasi: Fase ini digunakan untuk memperbaiki prototipe fungsional, terutama untuk memenuhi persyaratan non-fungsional.
5. Implementasi: Tahap implementasi mencakup penyerahan kepada pengguna yang dilanjutkan dengan penelaahan atas keberhasilan proyek.

2.2.8. Visualisasi Dashboard

Berikut adalah jenis-jenis dari visualisasi data yang dapat digunakan dalam dashboard [18]:

1. *Bar Chart*

Jenis visualisasi yang paling umum digunakan untuk membandingkan informasi dan menunjukkan data dengan perbedaan kuantitas, sehingga sangat efektif digunakan untuk data dengan jenis numerik.



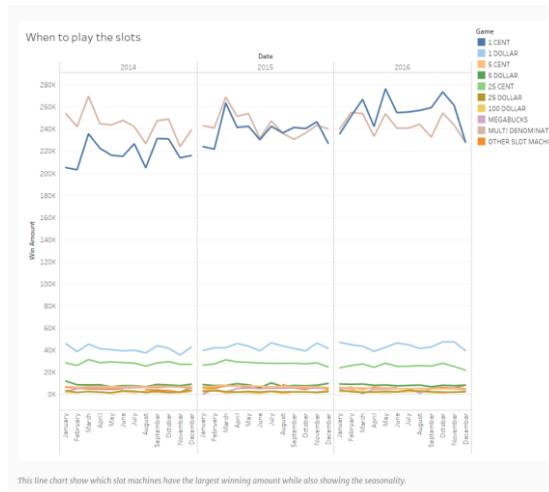
Gambar 2.1 Contoh *bar chart*

Untuk meningkatkan kualitas penggunaan bar chart pada dashboard, dapat dilakukan hal-hal sebagai berikut:

- Menampilkan beberapa bar chart dalam satu dashboard
- Memberikan warna untuk memperkuat kesan pada data
- Menggunakan stacked bars untuk data data yang berkaitan
- Mengombinasikan bar chart dengan peta sebagai filter

2. Line chart

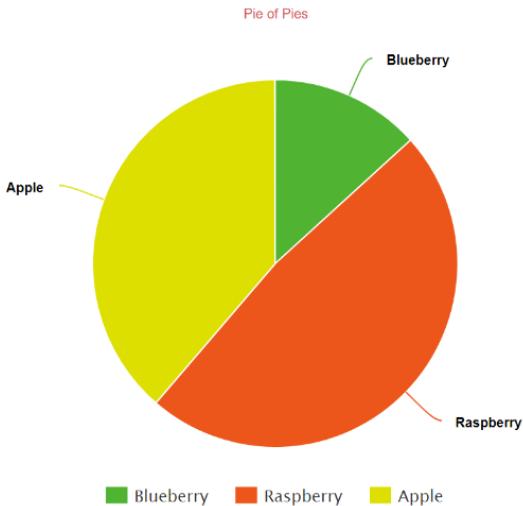
Line chart adalah grafik yang banyak digunakan untuk menghubungkan data-data dengan jenis angka, yang menghasilkan visualisasi berupa urutan nilai-nilai dalam data. Kegunaan utama dari chart ini adalah untuk menampilkan trend selama periode waktu tertentu. Gambar 2.2 menunjukkan contoh *line chart*.



Gambar 2.2 Contoh line chart

3. Pie Chart

Pie chart biasanya digunakan untuk menunjukkan proporsi atau prosentase dari sebuah informasi. *Pie chart* merupakan salah satu chart yang harus disandingkan dengan visualisasi lain untuk mendapat hasil yang lebih baik karena jika digunakan sebagai visualisasi utama hanya dapat menunjukkan proporsi dari data. Gambar 2.3 menunjukkan contoh *pie chart*.



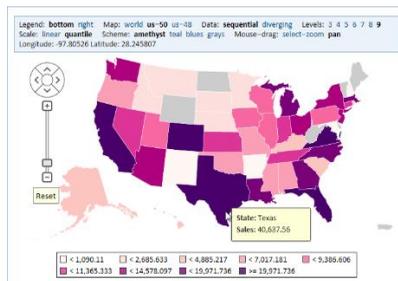
Gambar 2.3 Contoh *pie chart*

4. Map

Peta digunakan untuk melihat tampilan data berdasarkan lokasi. Untuk memaksimalkan penggunaan map, dapat dilakukan hal-hal berikut ini:

- Menggunakan map sebagai filter untuk grafik jenis lain
- Menggunakan *bubble map* di atas map

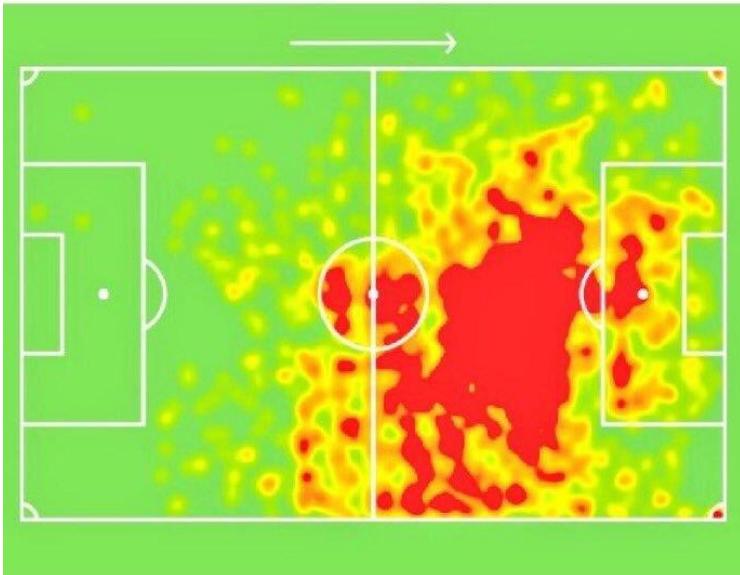
Gambar 2.4 menunjukkan contoh penggunaan map sebagai chart.



Gambar 2.4 contoh *map chart*

5. Heat Map

Heat maps adalah jenis visualisasi yang paling baik digunakan untuk membandingkan data dari dua kategori menggunakan warna, dan menunjukkan hubungan serta perpotongan dari kedua kategori tersebut. Gambar 2.5 merupakan contoh *heat map* yang menunjukkan rata-rata posisi Lionel messi ketika dia memenangkan *golden boot*.



Gambar 2.5 Contoh *heat map*

6. Highlight Table

Jenis visualisasi ini adalah perkembangan dari heat maps, dimana pada visualisasi ini daa juga ditampilkan dengan menggunakan warna, sehingga menampilkan informasi yang lebih mendetail terhadap sebuah informasi.

Population: Highlight table

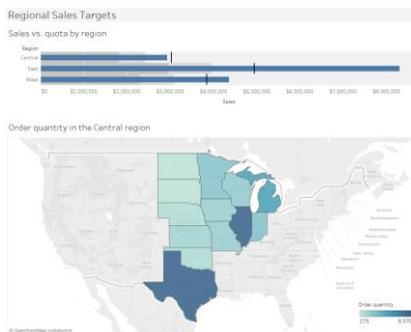
% Difference in Population - -04.77% + 06.03%

City	Year										
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
New York City	8,023,608	8,059,813	8,073,000	8,068,073	8,043,366	8,023,368	7,993,906	8,023,776	8,068,196	8,131,974	8,188,899
Los Angeles	3,702,062	3,733,328	3,762,432	3,784,279	3,792,660	3,788,197	3,767,960	3,751,872	3,763,666	3,781,938	3,797,462
Chicago	2,893,582	2,874,477	2,847,056	2,819,117	2,796,291	2,762,523	2,720,181	2,703,991	2,697,359	2,697,006	2,698,888
Houston	1,974,324	1,979,589	1,990,408	1,992,464	1,993,817	1,996,723	2,038,721	2,044,612	2,060,477	2,087,334	2,108,946
Philadelphia	1,514,563	1,505,465	1,498,493	1,493,802	1,492,882	1,490,861	1,488,710	1,493,309	1,499,731	1,514,694	1,528,306
Phoenix	1,326,682	1,342,908	1,357,196	1,368,806	1,386,946	1,421,720	1,441,241	1,450,901	1,466,729	1,482,879	1,499,911
San Antonio	1,153,843	1,161,680	1,176,718	1,188,981	1,205,860	1,223,976	1,246,274	1,270,187	1,291,232	1,310,846	1,334,466
San Diego	1,222,854	1,242,230	1,251,008	1,254,603	1,256,969	1,257,278	1,257,625	1,266,198	1,282,892	1,295,991	1,311,860
Dallas	1,188,168	1,196,974	1,190,681	1,181,976	1,174,231	1,171,519	1,177,916	1,179,667	1,189,805	1,192,663	1,202,739
Jacksonville	736,076	746,970	756,882	766,090	776,567	786,605	796,623	807,669	812,430	817,121	823,316
Indianapolis	782,760	787,039	787,420	789,466	791,038	793,606	796,247	802,751	808,013	815,982	821,793
San Francisco	777,885	780,885	772,723	766,160	761,825	763,016	768,650	778,673	781,162	801,922	805,463
Austin	676,776	691,119	698,644	690,293	696,973	705,423	705,873	709,641	704,946	700,972	707,533
Columbus	714,905	723,364	728,676	734,215	739,386	743,521	751,685	760,224	770,330	780,900	789,724
Fort Worth	547,484	566,402	585,542	602,759	619,117	637,139	660,948	689,763	705,526	727,267	744,723
Charlotte	577,676	591,176	601,869	613,674	626,629	645,100	669,217	691,668	705,993	724,333	734,443
Detroit	946,240	927,136	905,996	884,411	861,186	837,407	811,922	784,998	756,383	731,166	713,690
El Paso	568,224	572,621	576,719	582,199	590,713	597,634	609,650	616,846	626,920	639,017	652,129
Memphis	688,137	682,822	676,477	676,101	670,802	664,952	666,512	669,816	663,107	648,001	647,890
Baltimore	649,066	642,733	634,216	629,083	624,222	621,660	623,109	620,306	620,184	620,509	620,580
Boston	593,844	598,208	599,301	595,864	591,166	587,260	587,816	593,136	600,685	612,669	618,665
Seattle	563,434	567,169	566,946	566,711	568,719	573,336	580,609	586,966	593,588	604,189	610,728
Washington DC	572,046	574,604	573,158	568,502	567,754	567,136	570,681	574,404	580,236	592,228	604,453
Denver	555,651	561,976	566,790	562,588	560,796	561,691	566,895	564,395	576,721	589,011	604,434
Nashville	546,180	549,869	550,041	553,081	558,752	566,346	572,934	580,162	587,967	595,602	602,633
Milwaukee	596,382	596,396	596,181	594,283	592,008	588,665	586,395	586,446	587,858	592,180	595,588
Norland	529,845	535,765	540,833	542,496	536,636	538,053	544,669	555,306	562,860	576,104	585,802
Las Vegas	483,954	497,368	509,298	520,769	537,788	550,857	565,211	578,302	594,400	595,381	594,644
Oklahoma City	407,011	412,224	419,405	425,008	430,508	437,870	446,595	454,280	461,926	474,448	486,231
Albuquerque	443,161	447,627	448,551	478,028	489,868	502,924	514,723	524,690	531,461	539,216	547,563
Tucson	496,372	489,463	492,802	496,134	499,583	505,902	513,333	517,640	520,054	520,418	521,104
Fresno	430,801	436,525	444,360	453,763	462,601	466,179	471,748	477,639	484,443	490,262	496,153
Sacramento	406,593	421,149	431,196	439,790	446,049	449,523	452,338	455,216	459,341	463,468	467,580
Long Beach	461,978	464,856	467,330	468,897	468,781	467,090	465,445	460,328	460,543	461,782	462,813
Kansas City	442,421	443,846	445,289	445,623	445,928	447,015	449,652	452,322	454,348	457,897	460,852

Gambar 2.6 Contoh highlight table

7. Bullet chart

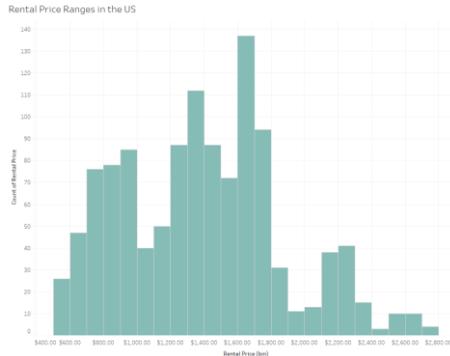
Bullet chart merupakan visualisasi yang menyajikan progress dari sebuah tujuan. Pada pokoknya, bullet chart merupakan pengembangan dari bar chart dengan menggunakan filter dan tujuan sebagai pembanding dengan progress. Gambar 2.7 merupakan contoh bullet chart



Gambar 2.7 Contoh bullet chart

8. Histogram *chart*

Histogram digunakan untuk melihat distribusi data di berbagai kategori, dan dapat pula digunakan jika ingin melihat keseimbangan data di berbagai kategori. Gambar 2.7 merupakan contoh penggunaan histogram.



Gambar 2.8 Contoh histogram

Untuk memaksimalkan penggunaan histogram, dapat dilakukan hal-hal sebagai berikut:

- Gunakan berbagai jenis kelompok data yang berbeda untuk menguji penggunaan histogram
- Gunakan filter

9. *Scatter plots*

Scatter plots ini digunakan untuk menganalisis hubungan antara variabel yang berbeda. *Scatter plots* dapat menunjukkan jika satu variabel dapat digunakan sebagai pemrediksi yang bagus untuk suatu data atau variabel tersebut berubah secara independen.

Sales and profit per department

Gambar 2.9 Contoh *scatter plots*

2.2.9. Tableau 9.2

Merupakan aplikasi pembuatan *dashboard* dan analisis data. Salah satu kehandalan dari tableau adalah kemampuannya untuk mengolah berbagai macam sumber data. Mulai dari basis data berbasis SQL seperti MySQL, SQL server, *postgresql* hingga anotasi seperti CSV dan Json.

Selain itu, aplikasi ini juga memungkinkan untuk visualisasi data berdasarkan lokasi dan waktu dan tanggal yang ekstensif.

2.2.10. Json

Json merupakan format data yang biasanya digunakan untuk melakukan pertukaran data tabular. Format data ini lebih ringan daripada format data yang lain [19]. Tipe data ini sangat mudah dipahami dan ditulis oleh manusia. Sementara untuk mesin juga mudah dalam mengolah dan membuat tipe data satu ini. Json merupakan bahasa pemrograman yang berbasis Javascript. Json merupakan bahasa berbasis teks yang independen namun

menggunakan konvensi yang familiar bagi *programmer* yang terbiasa dengan bahasa yang masuk dalam keluarga bahasa pemrograman C.

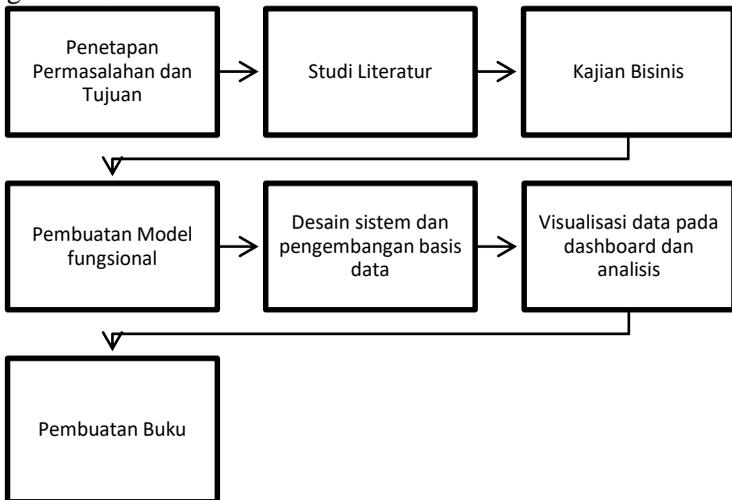
(halaman ini sengaja dikosongkan)

BAB III METODOLOGI PENELITIAN

Pada bab ini menjelaskan terkait metodologi yang akan digunakan sebagai panduan untuk menyelesaikan penelitian tugas akhir ini.

3.1. Metodologi Penelitian

Diagram Metodologi dari Tugas Akhir ini dapat dilihat pada gambar 3.1



Gambar 3.1 Metodologi penelitian

3.2. Tahap Pelaksanaan

Tahapan pelaksanaan berisi tahapan – tahapan yang dilaksanakan dalam penelitian tugas akhir ini.

3.2.1. Studi Literatur

Studi Pustaka didapatkan dari pengumpulan referensi diantaranya dari narasumber, buku dan penelitian sebelumnya. Studi pustaka dilakukan untuk lebih memahami dasar-dasar teori yang berhubungan dengan permasalahan. Studi pustaka dapat membantu peneliti dalam mencari sumber yang dapat

dipercaya dan latar belakang permasalahan yang diangkat dalam penelitian tugas akhir.

3.2.2. Kajian Bisnis

Tahapan yang dilakukan pada stage ini adalah melakukan kajian bisnis untuk mendapat data apa saja. Kajian bisnis dilakukan dengan observasi terhadap Instagram versi *web*. Observasi dilakukan dengan membuka halaman web instagram dan membuka elemen dari versi instagram tersebut. Hal ini dapat dilakukan dengan melihat sumber halaman (*page source*) instagram.

Selain itu dalam tahapan ini dilakukan bacaan terhadap penelitian sebelumnya. Hal ini bertujuan untuk membantu menentukan elemen apa dari halaman web instagram yang dapat dijadikan data.

3.2.3. Pembuatan Model Fungsional

Pada tahapan ini penulis melakukan pencarian terhadap *script* yang dapat mengambil data dari halaman web instagram. Hal ini dilakukan untuk memudahkan penulis dalam pengerjaan keseluruhan tugas akhir ini.

Setelah penulis mendapat aplikasi *scraping*, penulis melakukan modifikasi terhadap aplikasi tersebut sesuai dengan kebutuhan pengambilan data penulis. Yang didapat dari observasi elemen halaman *web* instagram. Aplikasi *crawling* ini harus dapat untuk mendapat data minimal *username*, jumlah *posts*, jumlah *followers*, jumlah *following*, tanggal *posting*, jumlah *likes* pada *post*, *user* yang berkomentar pada *post* tersebut dan isi komentarnya dan tanggal pengambilan data.

3.2.4. Desain Sistem dan Pembuatan Basis Data

Pada fase ini dilakukan desain sistem dan perancangan basis data. Sistem disini adalah mencakup keseluruhan *script* yang digunakan dalam pengerjaan tugas akhir ini. *Script* tersebut

antara lain, *script* pengambilan data dari halaman *web* instagram, *script* pembuatan basis data, *script* untuk memasukan data hasil pengambilan data halaman *web* instagram dan *script* analisis basis data.

Adapun pembuatan basis data sendiri disesuaikan dengan data yang berhasil didapat dari analisis halaman sumber instagram versi *web*. Pembuatan basis data yang pertama kali dilakukan adalah membuat skema basis data dan perancangan tabel pada basis data. Kemudian membuat basis data pada MySQL.

3.2.5. Analisis dan Visualisasi Data

Tahapan ini merupakan tahapan menampilkan data dari basis data. Setelah memasukan data pada basis data, hal berikutnya adalah membuat visualisasi data agar mudah dilakukan bacaan terhadap data tersebut. Setelah dapat divualisasikan hasil dapat dilakukan analisis dari data pada basis data. Beberapa analisis yang dilakukan antara lain membandingkan kondisi akun pada waktu tertentu dan membanding satu *post* pada waktu tertentu.

3.2.6. Penyusunan Laporan Tugas Akhir

Setelah seluruh proses dalam penelitian mengenai pengambilan data dan analisis terhadap instagram untuk menggali potensi pasar, selanjutnya melakukan penyusunan laporan tugas akhir.

3.3. Rangkuman Metodologi

Rangkuman metodologi berisi metodologi yang dilakukan dalam penelitian tugas akhir ini yang dimulai dari rangkaian aktifitas, tujuan, *input*, *output* dan metode yang digunakan seperti yang tertera pada tabel 3.1.

Table 3.1 Input dan Output Metodologi

<i>Input</i>	Proses	<i>Output</i>
Permasalahan	Identifikasi Masalah	Latar Belakang, Rumusan Masalah, Tujuan dan Manfaat
Latar Belakang, Rumusan Masalah, Tujuan dan Manfaat	Studi Pustaka	Dasar Teori dan Penelitian Sebelumnya
Halaman sumber instagram versi web	Kajian bisnis dan observasi	Elemen dan meta data pada instagram
Elemen dan meta data pada instagram	Pembuatan model fungsional	Data mentah instagram berupa json dan txt
Data mentah instagram berupa json dan txt	Desain system dan pengembangan basis data	<i>Script</i> dan basis data
<i>Script</i> dan basis data	Analisis dan visualisasi data	<i>Graph</i> dan tabel data beserta bacaanya
Elemen dan meta data pada instagram , Data mentah instagram berupa json dan txt , <i>Script</i> dan basis data, <i>Graph</i> dan tabel data beserta bacaanya	Penyusunan Tugas Akhir	Dokumen Tugas Akhir

BAB IV

PENGEMBANGAN APLIKASI *SCRAPING*

Pada bab ini akan membahas mengenai pengumpulan informasi dan data mentah dari halaman *web* instagram. Untuk mengumpulkan data, dilakukan observasi terhadap halaman *web* instagram dan hasil observasi tersebut digunakan untuk membuat *script scraping*.

4.1. Studi Kasus

Instagram merupakan salah satu media sosial terbesar saat ini. Dengan jutaan pengguna, instagram mempunyai cakupan yang luas, mulai dari olahraga sampai saluran penyampain berita terbaru.

Pada April tahun 2018, Facebook selaku pemilik instagram melakukan perubahan aturan yang membuat API instagram sudah tidak bisa lagi di liat oleh publik bahkan *programmer*. Facebook melakukan perubahan aturan ini untuk melindungi privasi dan data pengguna karena adanya kecurigaan dari publik Amerika bahwa Facebook melakukan jual beli data data pribadi pengguna facebook seperti hobi, pola pengguna dalam menerima informasi, produk apa yang dicari pengguna akhir akhir ini dan lain sebagainya. Maka sebagai langkah preventif, Instagram yang merupakan salah satu layanan andalan facebook menutup seluruh akses ke API mereka. Sehingga data yang sebelumnya mudah didapat menjadi lebih sulit untuk didapat.

Hal ini membuat rencana awal dari tugas akhir ini berubah. Pada akhirnya, pengambilan data dilakukan dengan membuka elemen halaman sumber instagram versi *web* dan secara manual mengambil data tersebut.

4.1.1. Observasi halaman *web* instagram

Dari halaman *web* instagram dapat dilihat beberapa informasi yang relevan digunakan untuk penelitian ini. Informasi yang relevan ini terdapat pada *script* yang menjadi sumber untuk

pembuatan aplikasi *crawling*. Ekspektasi sebelum membuka halaman sumber ini adalah beberapa data penting seperti *username*, jumlah *followers*, jumlah *posts* dan deskripsi akun portal berita dapat diperoleh dari halaman sumber instagram versi web.

```
<script type="application/ld+json">
{
  "@context": "http://schema.org",
  "@type": "Organization",
  "name": "detikcom",
  "alternateName": "@detikcom",
  "description": "Baca berita selengkapnya di sini",
  "url": "http://mtr.bio/detikcom/363693",
  "mainEntityOfPage": {
    "@type": "ProfilePage",
    "@id": "https://www.instagram.com/detikcom/"
  },
  "interactionStatistic": {
    "@type": "InteractionCounter",
    "interactionType": "http://schema.org/FollowAction",
    "userInteractionCount": "1149591"
  },
  "image": "https://www.instagram.com/static/images/ico/favicon-200.png/ab6eff595bb1.png",
  "email": "promosi@detik.com",
  "telephone": "+622179141177"
}
</script>
```

Gambar 4.1 Script json halaman utama

Pada gambar 4.1 ini menunjukkan *script json* yang berisi informasi mengenai akun instagram. Data ini diambil dari halaman sumber intagram versi web dari akun detikcom yang mempunyai url www.intagram.com/detikcom. Pada gambar 4.1 data data yang dapat diambil untuk penelitian ini adalah *name*, *alternateName*, *description*, *url*, *ProfilePage*, *userInteractionCount* dan *image*.

Data *name* merupakan data yang menunjukkan *username* dari pemilik akun. *Username* ini unik karena satu akun hanya punya satu *username*. *alternateName* merupakan alias yang digunakan dalam instagram untuk memudahkan dalam pencarian. *Description* merupakan informasi yang diberikan oleh pemilik akun instagarm untuk menjelaskan atau menerangkan mengenai diri mereka. *url* merupakan tautan keluar dari web instagram dan masuk ke halaman utama dari akun tersebut jika akun tersebut punya halaman internet sendiri. *ProfilePage* merupakan halaman utama instagram untuk akun tersebut. *UserInteractionCount* merupakan jumlah *followers* yang dimiliki oleh akun tersebut yang ditunjukkan pada halaman utama akun tersebut. *Image* merupakan metadata yang menunjukkan tautan gambar profile.

Table 4.1 Tipe data halaman utama *script* json

Data	Deskripsi
<i>name</i>	Tempat pengambilan data <i>username</i>
<i>alternateName</i>	Tempat pengambilan data alias
<i>description</i>	Tempat pengambilan data <i>description</i>
<i>url</i>	Tempat pengambilan data url
<i>ProfilePage</i>	Tempat pengambilan data tautan halaman profile
<i>userInteractionCount</i>	Tempat pengambilan data
<i>image</i>	Tempat pengambilan data tautan gambar profile

Pada tabel 4.1 menunjukkan tipe data dan deskripsi dari data tersebut untuk dilakukan pengambilan data pada *script* json pada halaman utama dari sebuah akun instagram.

Pada halaman utama terdapat beberapa data yang belum didapat seperti jumlah *post* yang dilakukan oleh akun tersebut. Terdapat beberapa data yang masih belum didapatkan pada gambar 4.1. Sehingga perlu dilakukan observasi lebih lanjut pada halaman utama. Observasi ini dilakukan untuk mendapat beberapa data seperti jumlah *post* dan jumlah *following* dari akun tersebut. Gambar 4.2 menunjukkan hasil lebih lanjut dari observasi halaman utama dari sebuah akun instagram. Pada gambar 4.2 terdapat beberapa metadata yang akan digunakan dalam membuat program *crawling* sederhana. Program *crawling* ini akan menyapu tiap elemen dalam halaman utama untuk mendapat data yang akan digunakan dalam penelitian ini.

```

<meta property="al:ios:app_name" content="Instagram" />
  <meta property="al:ios:app_store_id" content="389801252" />
  <meta property="al:ios:url"
content="instagram://user?username=detikcom" />
  <meta property="al:android:app_name" content="Instagram" />
  <meta property="al:android:package"
content="com.instagram.android" />
  <meta property="al:android:url"
content="https://www.instagram.com/_u/detikcom/" />

  <link rel="canonical" href="https://www.instagram.com/detikcom/" />
  <meta content="1.1m Followers, 15 Following, 3,231 Posts - See
Instagram photos and videos from @detikcom" name="description" />

  <meta property="og:type" content="profile" />
  <meta property="og:image" content="https://scontent-sin2-
1.cdninstagram.com/vp/20686bfaed35f2257ab6ac68c42c284d/5CC88480/t51.2885-
19/s150x150/36803558_2135119346776962_6245164936972992512_n.jpg?_nc_ht=sconte
nt-sin2-1.cdninstagram.com" />
  <meta property="og:title" content="@detikcom • Instagram photos
and videos" />
  <meta property="og:description" content="1.1m Followers, 15
Following, 3,231 Posts - See Instagram photos and videos from @detikcom" />
  <meta property="og:url"
content="https://www.instagram.com/detikcom/" />

```

Gambar 4.2 Sumber data halaman *followers*

Pada gambar 4.2 menunjukkan data apa saja yang ada pada halaman utama. Bisa terjadi redundansi data dengan *script* yang ada pada gambar 4.1 tetapi hal ini tidak menjadi masalah karena akan menggunakan salah satu *script* untuk mendapat satu jenis data.

Data data pada gambar 4.2 antara lain *username*, *url*, *followers*, *following*, *post* dan *description* serta gambar profil dari akun tersebut. Data pada *script* html yang ditunjukkan pada gambar 4.2 bisa dilihat pada tabel 4.2

Table 4.2 Tipe data halaman utama *script* html

Data	Deskripsi
<i>Username</i>	Metadata yang berisi <i>username</i> dari akun instagram

<i>url</i>	Metadata yang berisi tautan halaman profil dari akun instagram
<i>Followers</i>	Metadata yang berisi jumlah <i>followers</i> yang dimiliki oleh akun instagram
<i>Following</i>	Metadata yang berisi jumlah <i>following</i> yang dimiliki oleh akun instagram
<i>Post</i>	Metadata yang berisi jumlah <i>post</i> yang pernah dibuat oleh akun instagram
<i>Description</i>	Metadata yang berisi penjelasan mengenai akun instagram
<i>Image</i>	Metadata yang berisi tautan gambar profil dari akun instagram

Selain halaman profil, halaman lain yang perlu dilakukan analisis adalah halaman *posts*. Pada halaman *post* ada beberapa data yang menjadi data utama yang akan diambil. Data tersebut adalah tautan yang menunjukkan halaman tersebut, jumlah *likes*, jumlah komentar yang diberikan akun lain pada *post* tersebut, deskripsi atau isi dari *post* tersebut, lokasi dari gambar tersebut dan akun yang melakukan komentar pada akun tersebut dan isi dari komentar mereka. Halaman sumber yang dibuka pada gambar 4.3 merupakan halaman sumber dari sebuah *post* yang dibuat oleh akun instagram. Halaman *post* ini berbeda karena sudah mempunyai tautannya sendiri untuk membuka halaman ini.

```

<link rel="canonical"
href="https://www.instagram.com/detikcom/p/BsQDZAVlxWW/" />
<meta content="10.4k Likes, 586 Comments - @detikcom on Instagram: "Wakil
Direktorat Reserse Kriminal Khusus Polda Jatim, Ajun Komisaris Besar Polisi
Arman Asmara..." name="description" />
<meta property="og:site_name" content="Instagram" />
<meta property="og:title" content="@detikcom on Instagram: "Wakil
Direktorat Reserse Kriminal Khusus Polda Jatim, Ajun Komisaris Besar Polisi
Arman Asmara mengatakan penggerebekan terjadi karena..." />
<meta property="og:image" content="https://scontent-sin2-
1.cdninstagram.com/vp/569e8c5e487f08f4342654f7d1b9ela4/5CD9A0FF/t51.2885-
15/e35/47693294_245799842985418_1988373651163710287_n.jpg?_nc_ht=scontent-
sin2-1.cdninstagram.com" />
<meta property="og:description" content="10.4k Likes, 586 Comments -
@detikcom on Instagram: "Wakil Direktorat Reserse Kriminal Khusus Polda
Jatim, Ajun Komisaris Besar Polisi Arman Asmara..." />
<meta property="fb:app_id" content="124024574287414" />
<meta property="og:url"
content="https://www.instagram.com/p/BsQDZAVlxWW/" />
<meta property="instapp:owner_user_id" content="1527886079" />
<meta property="al:ios:app_name" content="Instagram" />
<meta property="al:ios:app_store_id" content="389801252" />
<meta property="al:ios:url"
content="instagram://media?id=1950073551140099478" />
<meta property="al:android:app_name" content="Instagram" />
<meta property="al:android:package" content="com.instagram.android" />
<meta property="al:android:url"
content="https://www.instagram.com/p/BsQDZAVlxWW/" />

```

Gambar 4.3 Sumber data halaman *post*

Dari gambar 4.3 beberapa data yang digunakan dalam penelitian ini adalah url, jumlah *likes*, jumlah komentar, tanggal, isi *post* dan akun yang berkomentar di *post* tersebut serta komentar mereka. Data tersebut ada pada *script* html instagram versi web.

Table 4.3 Data pada halaman *post*

Data	Deskripsi
url	Metadata yang berisi tautan halaman <i>post</i> yang dibuat oleh akun instaram

<i>Likes</i>	Metadata yang berisi jumlah <i>likes</i> yang dimiliki oleh <i>post</i> yang dibuat akun instagram
<i>Comments</i>	Metadata yang berisi jumlah komentar yang dibuat oleh akun instagram lain terhadap <i>post</i>
<i>Date</i>	Metadata yang berisi tanggal dibuatnya <i>post</i> oleh akun instagram
<i>Title</i>	Metadata yang berisi deskripsi dari <i>post</i> yang sudah dibuat oleh akun instagram
<i>User</i>	Metadata yang berisi <i>handle</i> dari akun instagram lain yang meninggalkan komentarnya pada <i>post</i>
<i>Description</i>	Metadata yang berisi komentar yang ditinggalkan oleh akun instagram lain pada <i>post</i>

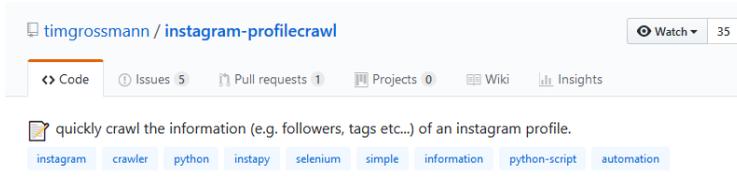
4.2. Pembuatan *Script Scraping*

Bagian ini berisikan pembuatan *script* untuk pengambilan data yang telah diidentifikasi pada observasi.

4.2.1. *Library* Python instagram-profilecrawl

Pada fase ini dilakukan pemilihan *library* yang akan digunakan untuk melakukan *scraping*. Dalam memilih *library* ini ada beberapa kriteria yang perlu diperhatikan yaitu penggunaan *python* dan *selenium* untuk memudahkan dalam otomasi halaman *web* instagram. Hal ini dilakukan untuk menginspeksi setiap elemen dari halaman *web* instagram.

Pada penelitian kali ini digunakan *library* dari github yang dibuat oleh pengguna bernama *timgrossmann*.



Gambar 4.4 Informasi *library* yang dipakai

Pada gambar 4.4 menunjukkan *library* yang akan digunakan dalam pengerjaan penelitian ini. *Library* ini memenuhi semua kebutuhan untuk *scraping* halaman *web* instagram. *Library* ini menggunakan python dan selenium untuk mengambil data dari elemen pada halaman sumber instagram versi *web*.

Pada *library* yang digunakan dalam penelitian ini, terdapat beberapa *dependencies* yang harus ada agar aplikasi dapat berjalan. *Dependencies* ini terdapat dalam keseluruhan sistem yang digunakan untuk *crawling* halaman instagram versi web.

Untuk lebih jelasnya mengenai *dependencies* ini dapat dilihat pada table 4.4. Pada table 4.4 dibuat daftar semua daftar *dependencies* yang harus ada dalam sistem *crawling* instagram versi web. Pada *library* ini terdapat beberapa *dependencies* yang menjadi kebutuhan utama berjalannya aplikasi. Terdapat beberapa *dependencies* yang harus didapatkan dari luar, dalam artian *dependencies* tersebut sudah ada sebelumnya dan tinggal melakukan instalasi. Seperti selenium yang merupakan komponen utama dalam otomasi sebuah halaman web. Selenium mempunyai webdriver yang bekerja dengan membuka halaman web menggunakan *browser* untuk dilakukan pencarian elemen. Aplikasi ini cukup populer dikalangan *tester* untuk menguji coba dan melihat bug dari halaman web yang telah dikembangkan.

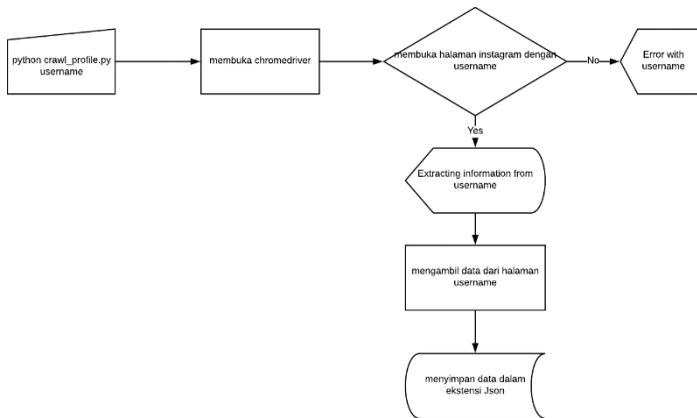
Table 4.4 *Dependencies sistem crawling*

<i>Dependencies</i>	Deskripsi
<i>Json</i>	Menyimpan hasil <i>crawling</i> dalam extensi Json
<i>Datetime</i>	Menentukan tanggal dan waktu pengambilan pada <i>post</i> dan waktu <i>scraped</i> atau pengambilan data
<i>Selenium</i>	Melakukan simulasi instagram versi web untuk pengambilan data
<i>Settings</i>	Pengaturan sistem pengambilan data seperti jumlah <i>post</i> yang diambil, waktu yang dibutuhkan untuk <i>crawling</i> antar <i>post</i>
<i>Datasaver</i>	Menyimpan data hasil <i>crawling</i>
<i>Extractor</i>	Menentukan data apa saja yang diambil dan struktur penyimpanan data

Dari *dependencies* dan hasil observasi terdapat beberapa data yang diambil. Data yang diambil adalah sebagai berikut seperti alias, *username*, *bio*, *prof_img*, jumlah *posts*, jumlah *followers*, jumlah *following*, *bio_url*, apakah akun tersebut merupakan akun pribadi dan *post*. *Post* tersebut berisikan data yang relevan untuk penelitian ini seperti tanggal, jumlah *likes*, url, img dan *preview_img*, komentar dan pengguna yang berkomentar pada *post* tersebut.

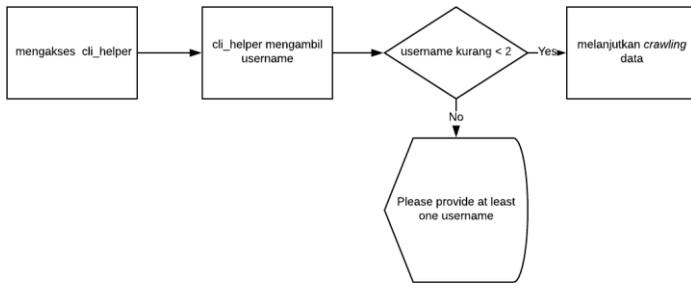
4.2.2. *Flowchart Script Library*

Untuk dapat mengambil data pada halaman instagram versi web digunakan *library* yang mempunyai *script* dengan logika aplikasi yang akan dijelaskan pada bagian ini. Terdapat beberapa *flowchart* pada aplikasi ini.



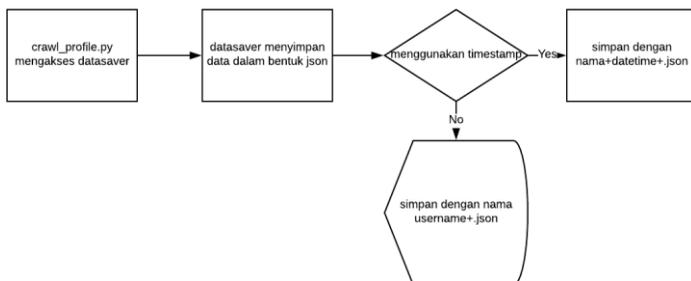
Gambar 4.5 *Flowchart crawl_profile.py*

Pada gambar 4.5 menjelaskan bagaimana aplikasi crawl bekerja. Langkah pertama adalah memasukan perintah python `crawl_profile.py username` dimana `username` disini adalah `username` dari akun instagram yang akan dilakukan pengambilan data. Setelah memasukan perintah tersebut, python akan mencari chromedriver dari selenium yang akan digunakan untuk mengotomasi halaman instagram pada browser chrome. Python akan membuka halaman instagram pada www.instagram.com/username jika terdapat pengguna dengan `username` tersebut maka pada bash akan menunjukkan pesan “*extracting information from username*” jika tidak ada halaman dengan `username` tersebut maka akan keluar pesan “*error with username*”. Setelah membuka halaman tersebut maka sistem akan mengambil data yang sesuai dengan *script extractor*. Data tersebut kemudian disimpan dalam bentuk json sesuai pada *script datasaver*.



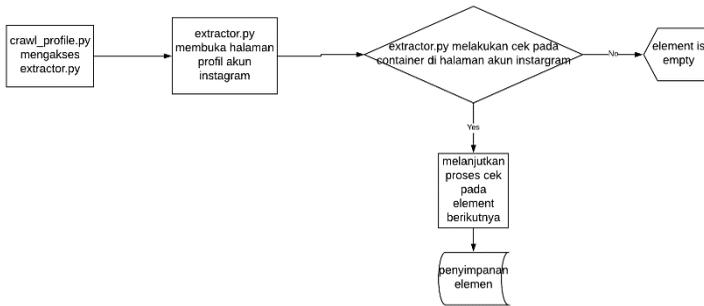
Gambar 4.6 Flowchart cli_helper.py

Gambar 4.6 menerangkan proses yang terjadi pada *dependencies* cli_helper pada aplikasi crawler. Fungsi utama dari script ini adalah untuk melakukan konfirmasi berapa *username* akun instagram yang dicari.



Gambar 4.7 Flowchart datasavers.py

Gambar 4.7 menunjukkan proses penyimpanan data hasil pengambilan data dalam bentuk ekstensi json. Script ini mempunyai fungsi utama untuk mengubah data hasil *crawling* menjadi berkas yang disimpan dalam bentuk json. Berkas json tersebut mempunyai penamaan *username+datetime.json* *username* menunjukkan akun yang telah di *crawling* dan *datetime* merupakan waktu selesai *crawling* yang Cuma berselisih beberapa detik dari waktu berkas json dibuat.



Gambar 4.8 Flowchart *extractor.py*

Gambar 4.8 merupakan *flowchart* sederhana dari *script extractor.py*. Secara prinsip *script* ini mengambil data dari setiap elemen yang ada pada halaman instagram versi web. Data itu diambil dengan melihat pada *container* yang ada pada halaman html instagram versi web. Data pada *container* antara lain terdapat pada tabel 4.6 menjelaskan mengenai data apa saja yang diambil dan dimana pengambilan datanya dilakukan pada instagram versi web.

Table 4.5 Pengambilan elemen halaman html instagram

Elemen	Keterangan
<i>Num_of_posts</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama kelas ('Y8-fY')
<i>Followers</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama kelas ('Y8-fY')
<i>Following</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama kelas ('Y8-fY')
<i>Prof_img</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama kelas ('-vDIg')

	dan elemen yang mempunyai nama tag ('h1')
<i>Bio</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama kelas ('-vDIg') dan elemen yang mempunyai nama tag ('span')
<i>Bio_url</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama kelas ('-yLUwa')
<i>Alias_name</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama kelas ('RR-M-') dan elemen yang mempunyai nama tag ('img') dengan atribut ('src')
<i>isprivate</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama kelas ('Nd_RI')
<i>Location</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama kelas ('M30cS') dan elemen yang mempunyai nama tag ('a')
<i>Img</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama tag ('img') dan atribut ('src')
<i>Date</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama kelas ('//a/time') dan atribut ('datetime')
<i>Likes count</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai xpath ('//article/div/section[2]//span')

<i>Comments count</i>	Diambil pada <i>container</i> dengan menghitung elemen yang mempunyai nama tag ('li')
<i>Comments</i>	Diambil pada <i>container</i> dengan elemen yang mempunyai nama tag ('ul')

4.2.3. Pengumpulan Data

Periode pengambilan data pada penelitian ini adalah pada tiga puluh Desember 2018 sampai dengan tiga Januari 2019. Pada periode itu tiap hari dilakukan pengumpulan data. Berkas json tersebut mempunyai struktur sebagaimana ditunjukkan oleh tabel 4.6.

Table 4.6 Struktur Data Json

Data	Deskripsi
<i>Alias</i>	Nama <i>handle</i> yang tertera pada halaman profil instagram
<i>Username</i>	Nama <i>handle</i> yang tertera pada halaman profil instagram
<i>Bio</i>	Deskripsi pemilik akun tersebut
<i>Prof_img</i>	Metadata yang menunjukkan gambar profil
<i>Num_of_posts</i>	Jumlah <i>post</i> yang dibuat oleh akun tersebut
<i>Followers</i>	Jumlah <i>followers</i> yang dimiliki akun tersebut
<i>Following</i>	Jumlah akun yang diikuti pemilik akun

<i>Bio_url</i>	Metadata yang menunjukkan url deskripsi akun
<i>Is_private</i>	Menunjukkan apakah akun tersebut akun pribadi atau terbuka untuk umum
<i>Caption</i>	Keterangan yang dibuat oleh pemilik akun terhadap gambar yang dia unggah
<i>Location</i>	Berisi keterangan lokasi gambar sesuai keinginan pemilik akun bisa tempat mengunggah ataupun tempat foto tersebut diambil
<i>Scraped</i>	Waktu pengambilan data
<i>Likes</i>	Jumlah <i>likes</i> yang dimiliki sebuah <i>post</i>
<i>URL</i>	Metadata yang menunjukkan url sebuah <i>post</i>
<i>Img</i>	Metadata yang menunjukkan tautan halaman gambar <i>post</i>
<i>Preview_img</i>	Metadata yang menunjukkan tautan halaman <i>preview</i> dari gambar <i>post</i>
<i>Comment</i>	Berisi komentar pada <i>post</i>
<i>Comments count</i>	Berisi jumlah komentar yang dibuat pemilik akun lain pada <i>post</i> tersebut.
<i>User</i>	Akun yang membuat komentar pada <i>post</i>

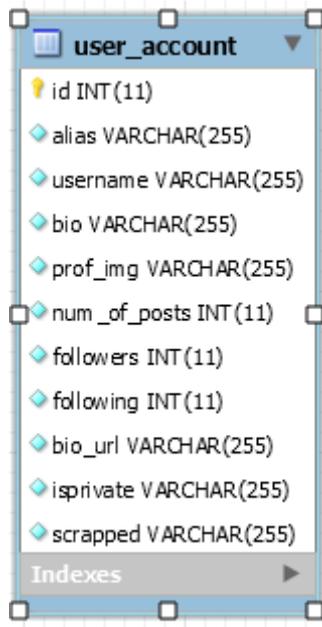
4.3. Desain Basis Data

Pada bagian ini akan menjelaskan mengenai desain dan pembuatan basis data. Dari json yang didapat setelah

melakukan *crawling* akan muncul data apa saja yang akan dipakai dalam basis data.

Langkah pertama dalam pengembangan basis data adalah dengan menganalisis data pada json dan menentukan data yang akan dipakai. Langkah berikutnya adalah membuat skema sesuai dengan data pada json dan tabel tabel yang diperlukan pada basis data. Dari skema tersebut kemudian dibuat tabel yang sesuai dengan data pada skema.

4.3.1. Pembuatan Skema Basis Data



Gambar 4.9 Skema tabel *user_account*

Gambar 4.9 menerangkan skema yang ada pada tabel *user_account* pada basis data crawl. Tabel *user_account* mempunyai sebelas kolom. Kolom *id* merupakan *primary key* dan *auto increment* yang berarti ketika ada data baru yang masuk pada tabel *user_account* kolom *id* akan mengisi dengan angka mulai dari satu, data baru yang masuk akan mempunyai

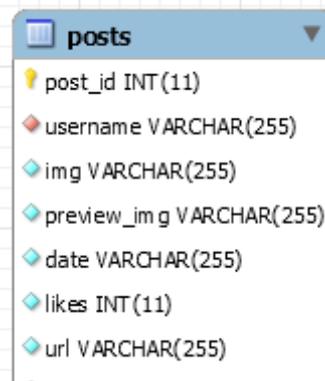
nilai dua dan data masuk berikutnya akan mempunyai nilai yang berurutan. Untuk lebih jelas mengenai kolom dalam tabel *user_account* pada basis data dapat dilihat pada tabel 4.7.

Table 4.7 Perancangan tabel *user_account*

Nama kolom	Tipe data	Deskripsi
<i>id</i>	INT(11)	<i>Primary key</i>
<i>alias</i>	VARCHAR(255)	Alias pemilik akun
<i>username</i>	VARCHAR(255)	Nama pemilik akun
<i>bio</i>	VARCHAR(255)	Deskripsi pemilik akun
<i>prof_img</i>	VARCHAR(255)	Gambar profile pemilik akun
<i>num_of_posts</i>	INT(11)	Jumlah <i>posts</i>
<i>followers</i>	INT(11)	Jumlah <i>followers</i>
<i>following</i>	INT(11)	Jumlah <i>following</i>
<i>bio_url</i>	VARCHAR(255)	Metadata tautan bio
<i>is_private</i>	VARCHAR(255)	Penanda akun pribadi
<i>scraped</i>	VARCHAR(255)	Tanggal pengambilan data

Pada tabel 4.7 menerangkan mengenai properti dari tabel *user_account* pada basis data *crawl*. Kolom *id* mempunyai tipe data integer dengan nilai sebelas, menggunakan tipe data integer karena data yang masuk adalah angka tiap kali ada data yang dimasukan akan mengikuti urutan memasukan data. Alias mempunyai tipe data VARCHAR dengan nilai 255 karena data yang akan dimasukan selain mempunyai huruf dalam *record* bisa juga mempunyai angka. Kolom *username* juga demikian, dalam data yang akan dimasukan dalam basis data

kemungkinan mempunyai huruf dan angka. Untuk kolom *bio*, *prof_img*, *bio_url*, *is_private* dan *scraped* juga mempunyai tipe data *varchar* karena bisa saja dalam *record* mempunyai huruf dan angka. Sementara *followers*, *following* dan *num_of_posts* mempunyai tipe data integer karena data yang akan diisikan kedalam basis data berupa angka.



Gambar 4.10 Skema tabel *posts*

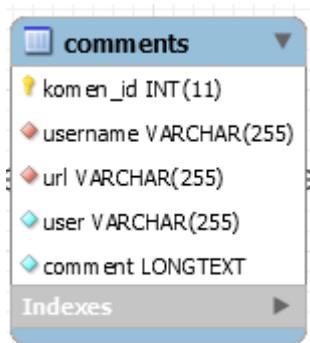
Gambar 4.10 merupakan skema yang menerangkan skema basis data pada tabel *posts* pada basis data *crawl*. Tabel *posts* mempunyai tujuh kolom. Kolom *post_id* merupakan *primary key* dan *auto increment* yang berarti ketika ada data baru yang masuk pada tabel *posts* kolom *post_id* akan mengisi dengan angka mulai dari satu, data baru yang masuk akan mempunyai nilai dua dan data masuk berikutnya akan mempunyai nilai yang berurutan. *Username* merupakan *foreign key* yang mempunyai referensi ke tabel *user_account*.

Table 4.8 Perancangan tabel *posts*

Nama kolom	Tipe data	Deskripsi
<i>post-id</i>	INT(11)	<i>Primary key</i>
<i>username</i>	VARCHAR(255)	<i>Foreign key</i> kategori <i>user_account</i>

img	VARCHAR(255)	Metadata tautan gambar
preview_img	VARCHAR(255)	Metadata <i>preview</i> gambar
date	VARCHAR(255)	Tanggal <i>posting</i>
<i>likes</i>	INT(11)	Jumlah <i>likes</i>
url	VARCHAR(255)	Metadata tautan <i>post</i>

Pada tabel 4.11 menerangkan properti dari tabel *posts* pada basis data *crawl*. Kolom *post_id* mempunyai tipe data integer karena begitu ada data yang masuk kedalam database akan dibuat nilai yang berurutan sesuai urutan masuk dimulai dari satu. Karena dari json data yang berkorespondensi dengan kolom *likes* pada tabel *posts* mempunyai nilai angka. Sementara untuk kolom *username*, *img*, *preview_img*, *date* dan *url* mempunyai tipe data *varchar* karena ada kemungkinan data yang berasal dari json dapat berupa kombinasi angka dan huruf.



Gambar 4.11 Skema tabel *comments*

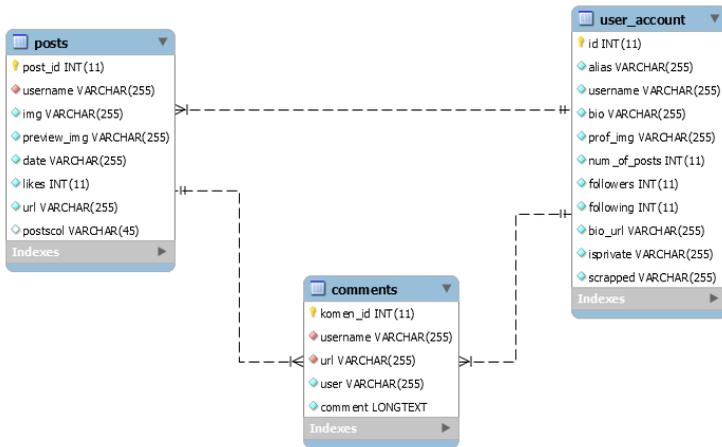
Pada gambar 4.11 menerangkan mengenai skema dari tabel *comments* pada basis data *crawl*. Tabel *comments* mempunyai lima kolom. Kolom *id* merupakan *primary key* dan *auto increment* yang berarti ketika ada data baru yang masuk pada tabel *user_account* kolom *id* akan mengisi dengan angka mulai

dari satu, data baru yang masuk akan mempunyai nilai dua dan data masuk berikutnya akan mempunyai nilai yang berurutan. Pada tabel *comments* mempunyai dua *foreign key* salah satunya berasal dari tabel *user_account* dan satunya lagi berasal dari tabel *posts*. *Username* merupakan kolom yang berasal dari tabel *user_account* dan *url* merupakan kolom yang berasal dari tabel *posts*.

Table 4.9 Perancangan tabel *comments*

Nama kolom	Tipe data	Deskripsi
<i>komen_id</i>	INT(11)	<i>Primary key</i>
<i>username</i>	VARCHAR(255)	<i>Foreign key</i> kategori <i>user_account</i>
<i>url</i>	VARCHAR(255)	<i>Foreign key</i> kategori <i>posts</i>
<i>user</i>	VARCHAR(255)	Akun instagram lain yang melakukan komentar pada <i>post</i>
<i>comment</i>	LONGTEXT	Isi komentar

Tabel 4.8 menerangkan mengenai properti dari tabel *comments* pada basis data *crawl*. Kolom *komen_id* mempunyai tipe data integer karena data yang masuk adalah *record* angka yang dimulai dari satu. Kolom *username*, *url* dan *user* memungkinkan adanya kombinasi antara huruf dan angka pada nilainya sehingga digunakan tipe data VARCHAR. Sementara untuk *comment* karena kemungkinan ada nilai yang panjang dari kolom ini maka menggunakan tipe data LONGTEXT.

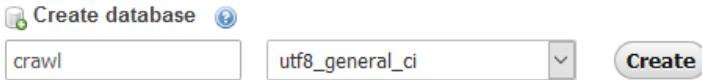


Gambar 4.12 Setiap tabel pada basis data

Gambar 4.12 berisikan setiap tabel pada basis data *crawl*. Basis data *crawl* mempunyai tiga tabel yang digunakan untuk menyimpan data json hasil *crawling*. Tabel pertama adalah tabel *user_account* yang berisikan informasi umum dari sebuah akun instagram. Sementara tabel *posts* merupakan tabel yang berisi informasi *post* yang dibuat oleh akun instagram sehingga mempunyai *username* yang berasal dari tabel *user_account*. Tabel *comments* merupakan tabel yang berisi informasi mengenai komentar yang dibuat oleh pengguna instagram terhadap pembuat *posts* dan isi komentar yang dibuat pengguna instagram. Karena hal itulah *comments* mempunyai kolom yang berasal dari tabel *user_account* dan tabel *posts*. Kolom yang berasal dari tabel *user_account* adalah kolom *username* dan kolom yang berasal dari tabel *posts* adalah kolom *url*.

4.3.2. Pembuatan Basis Data

Pada bagian ini akan menjelaskan bagaimana proses pembuatan basis data yang digunakan dalam penelitian kali ini. Gambar 4.13 menunjukkan pembuatan nama basis data dan *collation* yang digunakan.



Gambar 4.13 Pembuatan basis data *crawl*

Basis data yang akan dibuat adalah basis data *crawl* yang memiliki tiga tabel utama yaitu *user_account*, *posts* dan *comment*. Basis data ini mempunyai tipe *collation* *utf8_general_ci*. Penggunaan tipe *collation* ini adalah agar memudahkan pengisian database dari json. Karena pada berkas json juga digunakan *utf8* sebagai karakter *encoding*.

```
CREATE TABLE `crawl`.`user_account` ( `id` INT(11) UNSIGNED NOT
NULL AUTO INCREMENT , `alias` VARCHAR(255) NOT NULL , `username`
VARCHAR(255) NOT NULL , `bio` VARCHAR(255) NOT NULL , `prof_img`
VARCHAR(255) NOT NULL , `num_of_posts` INT(11) NOT NULL ,
`followers` INT(11) NOT NULL , `following` INT(11) NOT NULL ,
`bio_url` VARCHAR(255) NOT NULL , `isprivate` VARCHAR(255) NOT
NULL , `scrapped` VARCHAR(255) NOT NULL , PRIMARY KEY (`id`))
ENGINE = MyISAM;
```

Gambar 4.14 Script pembuatan tabel *user_account*

Gambar 4.14 menyajikan *script* SQL pembuatan tabel *user_account*. Tabel *user_account* berisi *id* sebagai *primary key*, *alias*, *username*, *bio*, *prof_img*, *num_of_posts*, *followers*, *following*, *bio_url*, *isprivate* dan *scraped* yang merupakan tanggal pengambilan data.

```
CREATE TABLE `crawl`.`posts` ( `post_id` INT(11) UNSIGNED NOT
NULL AUTO INCREMENT , `username` VARCHAR(255) NOT NULL , `img`
VARCHAR(255) NOT NULL , `preview_img` VARCHAR(255) NOT NULL ,
`date` VARCHAR(255) NOT NULL , `likes` INT(11) NOT NULL , `url`
VARCHAR(255) NOT NULL , PRIMARY KEY (`post_id`)) ENGINE = MyISAM;
```

Gambar 4.15 Script pembuatan tabel *posts*

Gambar 4.15 berisi *script* SQL pembuatan tabel *posts*. Tabel *posts* berisi *post_id*, *username*, *img*, *preview_img*, *date*, *likes* dan *url*. *Post_id* digunakan sebagai *primary id*, karena ada akan ada beberapa *post* yang merupakan duplikat dari *post* lain yang

diambil pada waktu yang berbeda. *Username* merupakan *foreign key* yang diambil dari *user_account*.

#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra	Action
1	post_id	int(11)		UNSIGNED	No	None		AUTO_INCREMENT	Change Drop Primary Unique Index More
2	username	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
3	img	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
4	preview_img	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
5	date	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
6	likes	int(11)			No	None			Change Drop Primary Unique Index More
7	url	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More

Check all With selected: Browse Change Drop Primary Unique Index Fulltext

Gambar 4.16 Struktur tabel *posts*

Gambar 4.16 merupakan struktur dari basis data *posts* sebelum kolom *username* diubah menjadi *foreign key* yang berdasarkan dari tabel *user_account*.

```

1 ALTER TABLE posts
2 ADD CONSTRAINT fk_username
3 FOREIGN KEY (username) REFERENCES user_account(username)

```

Gambar 4.17 Menambah *foreign key username*

Gambar 4.17 merupakan *script* yang digunakan untuk mengubah *username* pada tabel *posts* menjadi *foreign key* yang berdasarkan pada *username* pada tabel *user_account*.

#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra	Action
1	post_id	int(11)		UNSIGNED	No	None		AUTO_INCREMENT	Change Drop Primary Unique Index More
2	username	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
3	img	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
4	preview_img	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
5	date	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
6	likes	int(11)			No	None			Change Drop Primary Unique Index More
7	url	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More

Check all With selected: Browse Change Drop Primary Unique Index Fulltext

1 column(s) after url

Action	Keyname	Type	Unique	Packed	Column	Cardinality	Collation	Null	Comment
Edit	PRIMARY	BTREE	Yes	No	post_id	0	A	No	
Edit	fk_username	BTREE	No	No	username	A		No	

Gambar 4.18 Struktur tabel terbaru

Gambar 4.18 merupakan struktur tabel *posts* terbaru yang telah diperbaharui dengan mengubah *username* menjadi *foreign key*.

```
CREATE TABLE `crawl`.`comments` ( `komen_id` INT(11) UNSIGNED NOT
NULL AUTO INCREMENT , `username` VARCHAR(255) NOT NULL , `url`
VARCHAR(255) NOT NULL , `user` VARCHAR(255) NOT NULL , `comment`
LONGTEXT NOT NULL , PRIMARY KEY(`komen_id`)) ENGINE = MyISAM;
```

Gambar 4.19 Script pembuatan tabel *comments*

Gambar 4.19 merupakan *script* SQL yang digunakan untuk membuat tabel *comments*. Tabel *comments* berisi *komen_id*, *username*, *url*, *user* dan *comment*. *Komen_id* digunakan sebagai *primary key* agar memudahkan *indexing* dan ada data yang sama tetapi diambil pada waktu yang berbeda. *Username* merupakan *foreign key* yang berasal dari *user_account* sementara *url* merupakan *foreign key* dari *posts*.

#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra	Action
<input type="checkbox"/>	1	komen_id	int(11)	unsigned	No	None	AUTO_INCREMENT		Change Drop Primary Unique Index More
<input type="checkbox"/>	2	username	varchar(255)	utf8_general_ci	No	None			Change Drop Primary Unique Index More
<input type="checkbox"/>	3	url	varchar(255)	utf8_general_ci	No	None			Change Drop Primary Unique Index More
<input type="checkbox"/>	4	user	varchar(255)	utf8_general_ci	No	None			Change Drop Primary Unique Index More
<input type="checkbox"/>	5	comment	longtext	utf8_general_ci	No	None			Change Drop Primary Unique Index More

↑ Check all With selected: Browse Change Drop Primary Unique Index Fulltext

Print Propose table structure Move columns Improve table structure

Add 1 column(s) after comment Go

Action	Keyname	Type	Unique	Packed	Column	Cardinality	Collation	Null	Comment
Edit Drop	PRIMARY	BTREE	Yes	No	komen_id	0	A	No	

Gambar 4.20 Struktur tabel awal *comments*

Gambar 4.20 merupakan struktur awal tabel *comments* sebelum ditambahkan *foreign key* *username* dan *url*.

```
1 ALTER TABLE comments
2 ADD CONSTRAINT fk_username
3 FOREIGN KEY (username) REFERENCES user_account (username)
4 |
```

Gambar 4.21 Script penambahan *fk username*

Gambar 4.21 merupakan penambahan *foreign key username* pada tabel. *Username* ini didapat merupakan kolom yang berasal dari *user_account* tetapi juga ada di tabel *posts* juga. *Username* pada tabel ini mempunyai *constraint* *fk_username*.

```

1 ALTER TABLE comments
2 ADD CONSTRAINT fk_post
3 FOREIGN KEY (url) REFERENCES posts(url)

```

Gambar 4.22 Script penambahan *fk url*

Gambar 4.22 merupakan *script* yang digunakan untuk membuat *foreign key url* yang mereferensi dari tabel *posts*. *Foreign key url* sendiri pada tabel *comments* mempunyai *constraint fk_post*.

#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra	Action
1	komen_id	int(11)		UNSIGNED	No	None		AUTO_INCREMENT	Change Drop Primary Unique Index More
2	username	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
3	url	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
4	user	varchar(255)	utf8_general_ci		No	None			Change Drop Primary Unique Index More
5	comment	longtext	utf8_general_ci		No	None			Change Drop Primary Unique Index More

Action	Keyname	Type	Unique	Packed	Column	Cardinality	Collation	Null	Comment
Edit Drop	PRIMARY	BTREE	Yes	No	komen_id	0	A	No	
Edit Drop	fk_username	BTREE	No	No	username	A	No	No	
Edit Drop	fk_post	BTREE	No	No	url	A	No	No	

Gambar 4.23 Struktur akhir tabel *comments*

Gambar 4.23 merupakan struktur akhir tabel *comments* yang sudah berisi dua *foreign key* baru yaitu *username* dan *url*.

Table	Action	Rows	Type	Collation	Size	Overhead
comments	Empty Drop	0	MyISAM	utf8_general_ci	4 KLB	-
posts	Empty Drop	0	MyISAM	utf8_general_ci	4 KLB	-
user_account	Empty Drop	0	MyISAM	utf8_general_ci	1 KLB	-
3 tables	Sum	0	MyISAM	utf8_general_ci	9 KLB	0 B

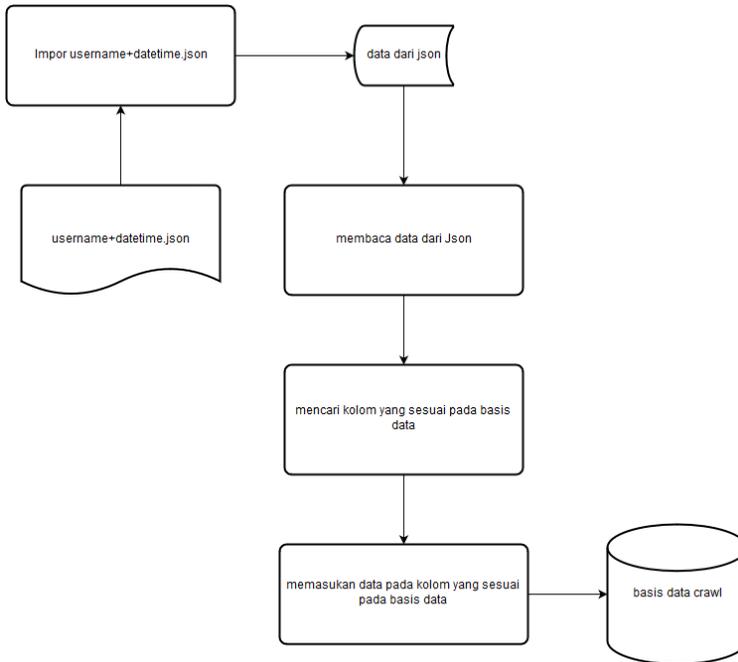
Gambar 4.24 Basis data *crawl*

Gambar 4.24 merupakan tabel yang ada pada basis data *crawl* beserta propertiinya.

4.3.3. Pengisian Basis Data

Pengisian basis data dilakukan dengan membuat tiga *script* yang akan memasukan data ke dalam basis data dari berkas *json*. *Script* yang digunakan merupakan *script php* yang mencari isi *array* dari kata kunci yang sudah disediakan pada *script*. Berikut ini merupakan *script* untuk mengambil data untuk tabel *comments*. Walaupun ketiga *script* pada modul memasukan data dari *json* kedalam basis data berbeda dari segi data yang

dimasukan ke dalam basis data. Tetapi secara tahap ketiga *script* tersebut mempunyai prinsip yang sama.

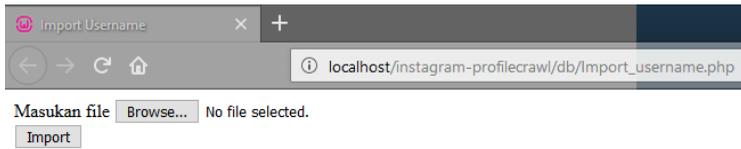


Gambar 4.25 Flowchart memasukan data

Pada gambar 4.25 untuk memasukan data dari Json ke basis data perlu dilakukan perubahan data dari *string* menjadi *array*. Untuk membaca data json tersebut terlebih dahulu dilakukan import menggunakan halaman web sederhana. Setelah data tersebut diubah menjadi *array* maka langkah selanjutnya adalah memasukan data sesuai dengan nama pada json dan tabel pada basis data. Setelah dicocokkan antara dua elemen, data yang akan dimasukan disiapkan terlebih dahulu dalam *statement* pada php. Setelahnya data dimasukan pada database.

Untuk setiap data dari json dimasukan kedalam tabel tersendiri sesuai dengan basis data yang telah dibuat seperti data *username*, *url*, *user* dan *comment* untuk tabel *comments* pada basis data. Data *username*, *img*, *preview_img*, *date*, *likes* dan

url pada tabel *posts*. Data yang akan dimasukkan kedalam tabel antara lain alias, *username*, *bio*, *prof_img*, *num_of_posts*, *followers*, *following*, *bio_url*, *isprivate* dan *scraped*.



Gambar 4.26 Tampilan memasukan data json ke basis data

Gambar 4.26 merupakan tampilan dari *script* sederhana yang digunakan untuk memasukan data dari json ke basis data. Pada tampilan halaman web tersebut mempunyai tombol *browse* dan *import* yang akan digunakan untuk memindahkan data dari json ke database.

(halaman ini sengaja dikosongkan)

BAB V UJI COBA

Bab ini menjelaskan mengenai desain dan memasukan basis data untuk penelitian kali ini. Dari json data yang telah diperoleh dapat dibagi menjadi tiga tabel yaitu *username*, *posts* dan komentar

5.1. Uji Coba

Pada bagian ini menjelaskan mengenai pengujian aplikasi *crawling* dan memasukan data ke database. Hasil test case dapat dilihat pada tabel dibawah ini.

Table 5.1 Uji Coba

<i>User story</i>	Deskripsi	Scenario	Status
<i>User</i> dapat melakukan <i>crawling</i> pada halaman akun instagram versi web	<i>Crawling</i>	<i>User</i> mengetikan perintah python <i>crawl_profile.py username</i> (yang akan dicrawl dan terdaftar di instagram) pada bash	Berhasil
		<i>User</i> mengetikan perintah python <i>crawl_profile.py username</i> (yang akan dicrawl dan tidak di instagram) pada bash	Gagal
<i>User</i> dapat memasukan data	Database	<i>User</i> memasukan berkas json	Berhasil

<i>user_account</i> pada tabel <i>user_account</i>		dengan struktur yang sesuai	
		<i>User</i> memasukan berkas json dengan struktur yang tidak sesuai	Gagal
<i>User</i> dapat memasukan data <i>posts</i> pada tabel <i>posts</i>	Database	<i>User</i> memasukan berkas json dengan struktur yang sesuai	Berhasil
		<i>User</i> memasukan berkas json dengan struktur yang tidak sesuai	Gagal
<i>User</i> dapat memasukan data komentar pada tabel <i>comments</i>	Database	<i>User</i> memasukan berkas json dengan struktur yang sesuai	Berhasil
		<i>User</i> memasukan berkas json dengan struktur yang tidak sesuai	Gagal

5.2. Spesifikasi

Pada bagian ini membahas mengenai kebutuhan perangkat lunak dan perangkat keras dalam pengerjaan tugas akhir ini.

5.2.1. Perangkat Keras

Requirement

CPU: minimal 2 Core

HD: 25GB

Memori: 512MB RAM

Existing

CPU: Intel(R) Core(TM) i3-2120 CPU @ 3.30GHz (4 CPUs)

HD: 1TB

Memori: 1932MB RAM

5.2.2. Perangkat Lunak

Requirement

Browser: Mozilla 40.0 , Google Chrome, Microsoft Edge

OS: Linux, Unix-like, OS X, Windows, tableau 9.2, python, wamp server

Existing

Browser: Mozilla 43.0.1

OS: Windows XP Professional (5.1, Build 2600) Service Pack 2 (2600.xpsp_sp2_rtm.040803-2158), tableau 9.2, python, wamp server

(halaman ini sengaja dikosongkan)

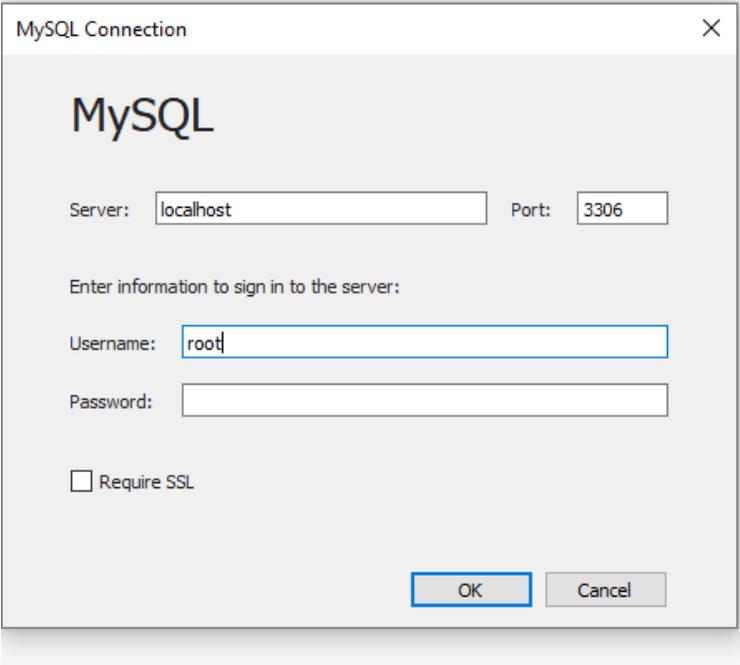
BAB VI VISUALISASI DAN ANALISIS DATA

Bab ini menjelaskan visualisasi dan dilakukan analisis terhadap visualisasi tersebut. Digunakan tableau versi 9.2 untuk melakukan visualisasi dan analisis data.

6.1. Pengolahan Data Pada *Tableau 9.2*

Tableau 9.2 digunakan untuk memudahkan dalam pembacaan data dan visualisasinya.

6.1.1. Memasukan Data Pada Tableau 9.2

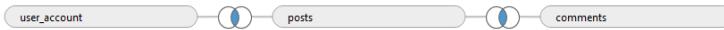


The image shows a 'MySQL Connection' dialog box. At the top, it says 'MySQL Connection' with a close button. Below that is the 'MySQL' logo. There are two input fields: 'Server:' with 'localhost' and 'Port:' with '3306'. Below these is the instruction 'Enter information to sign in to the server:'. There are two more input fields: 'Username:' with 'root' and 'Password:'. At the bottom left, there is a checkbox labeled 'Require SSL' which is unchecked. At the bottom right, there are two buttons: 'OK' and 'Cancel'.

Gambar 6.1 Koneksi ke server basis data

Pada 6.1 diatas menjelaskan mengenai bagaimana masuk pada basis data penyimpanan data hasil *crawl* halaman instagram versi web. Server yang digunakan adalah localhost dengan port 3306 dan *username* root dengan tidak ada password. Localhost

merupakan server lokal dimana basis data *crawl* ada. Server tersebut dapat diakses pada port 3306. Untuk dapat mengakses basis data tersebut dibutuhkan *username* dan *password*, secara *default* pada MySQL *username* untuk mengakses server adalah *root* dan *password* yang digunakan untuk mengakses database tersebut adalah *field* dikosongkan.



Gambar 6.2 Potongan tiap tabel pada tableau

Gambar 6.2 merupakan perpotongan antar tabel pada tableau 9.2 dari data yang berasal dari MySQL. Pada tabel *user_account* terdapat *username* yang juga dimiliki oleh tabel *posts* dan tabel *comments*.

Setelah masuk pada server basis data, basis data yang digunakan adalah *crawl*. Data yang digunakan pada penelitian ini sendiri punya potongan seperti diatas. *Username* pada *post* dan *comment* sama dengan *username* pada *user_account* (*foreign key*). Sementara tabel *posts* dan *comments* berpotongan pada kolom url.

ID	Username	URL	Comment	Post ID	Username (Post)	Img	Preview Img	Date	Like	Url
1	delikcom	https://www.instagram.com/delikcom	"Halo guys aku baru ..."	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
2	delikcom	https://www.instagram.com/pastorrahayuh	"47'88'NY'WAJDU"	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
3	delikcom	https://www.instagram.com/pasangefisudak...	"Mau bu, pak, mbak, m...	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
4	delikcom	https://www.instagram.com/afyulid	"JASA PENJAJAH FOL..."	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
5	delikcom	https://www.instagram.com/pambukamisan1	@pasangefisudak...	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
6	delikcom	https://www.instagram.com/eruan_pafin_ow	"Vah Gagah"	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
7	delikcom	https://www.instagram.com/pasangefisudak...	@pambukamisan1 ay...	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
8	delikcom	https://www.instagram.com/kaprian	"Prof @dikorlatamam..."	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
9	delikcom	https://www.instagram.com/tekniamy	"Kak bisa ngomong beg..."	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
10	delikcom	https://www.instagram.com/duff99	"Kaga masalah klo ga..."	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
11	delikcom	https://www.instagram.com/yugentrip	"Senen gratis loh guys..."	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram
12	delikcom	https://www.instagram.com/abegpurnomo	"Munduk apa sampan k..."	1	delikcom	https://content.s-i2-1...	https://content.s-i2-1...	2018-12-28T04:26:48.00...	30,934	https://www.instagram

Gambar 6.3 Data yang dianalisis pada tableau

Hasilnya adalah data seperti pada gambar 6.3. Setelah keluar tabel hasil bacaan dari basis data *crawl*. Dilakukan ekstraksi data agar memudahkan dalam visualisasi dan analisis data karena tidak perlu untuk membuka basis data lagi dalam pembuatan visualiasi data. Kolom yang terdapat pada tabel bacaan untuk tableau 9.2 antara lain kolom *username*, *bio*, *prof_img*, *num_of_posts*, *followers*, *following*, *bio_url*, *isprivate*, *scraped*, *user*, *comment*, *img*, *preview_img*, *date*, *likes* dan *url*.

6.1.2. Visualisasi Data

The screenshot shows a Tableau worksheet with the following configuration:

- Columns:** Measure Names
- Rows:** Username, Scrapped
- Filters:** Measure Names
- Marks:** Abc Automatic (dropdown), Color, Size, Text (all disabled), Detail, Tooltip (both disabled), Measure Values (dropdown)
- Measure Values:** AVG(Followers), AVG(Num Of Posts)

Username	Scrapped	Avg. Followers	Avg. Num Of Posts
detickcom	2018-12-31 16:26:10	1,135,647	3,192
	2019-01-01 16:27:10	1,137,620	3,208
	2019-01-02 18:36:56	1,140,277	3,213
	2019-01-03 12:02:16	1,141,919	3,215
kompascom	2018-12-31 16:35:35	640,033	2,273
	2019-01-01 16:37:10	640,848	2,274
	2019-01-02 18:46:36	641,950	2,280
	2019-01-03 12:12:02	642,702	2,280
sindonews	2018-12-31 16:39:03	61,584	1,879
	2019-01-01 16:40:26	61,839	1,881
	2019-01-02 18:50:22	62,181	1,882
	2019-01-03 12:15:20	62,452	1,884
tribunnews	2018-12-31 17:03:21	597,861	4,995
	2019-01-01 17:06:46	598,897	5,001
	2019-01-02 19:11:38	600,342	5,009
	2019-01-03 12:39:24	601,225	5,015

Gambar 6.4 Perbandingan jumlah *followers* dan *posts*

Pada gambar 6.4 menunjukkan perubahan jumlah *followers* dan jumlah *posts* pada empat waktu pengambilan data untuk empat *username* yaitu *detickcom*, *kompascom*, *sindonews* dan *tribunnews*. Gambar 6.4 belum dilakukan perubahan dalam visualisasi data, hanya menampilkan tabel tanpa ada perubahan yang berarti. Data tersebut diurutkan berdasarkan waktu pengambilan data (*scrapped* disini adalah waktu pengambilan data). Untuk lebih jelasnya bagaimana perubahan jumlah *followers* dan jumlah *post* dapat dianalisis lebih jauh per tiap *username*.

Pada gambar 6.4 paling kiri merupakan *username* dari akun yang diambil datanya. Kanan dari *username* adalah *scrapped* yaitu waktu pengambilan data, *scrapped* mempunyai format tahun-bulan-tanggal jam-menit-detik. Kanan dari kolom *scrapped* adalah kolom *avg. followers*, kolom ini menunjukkan jumlah *followers* yang dimiliki oleh *username* pada waktu pengambilan data. Avg. atau rata rata disini digunakan untuk

menormalisasi jumlah *followers*. Hal ini perlu dilakukan karena kolom *followers* terisi berkali kali pada *record comments*. Sama halnya dengan *num of posts*, diperlukan penggunaan rata rata untuk menormalisasi jumlah *posts* yang dibuat. Kolom *num of posts* sendiri adalah kolom yang menunjukkan aktifitas akun tersebut karena *num of posts* adalah jumlah *posts* yang dibuat akun tersebut.

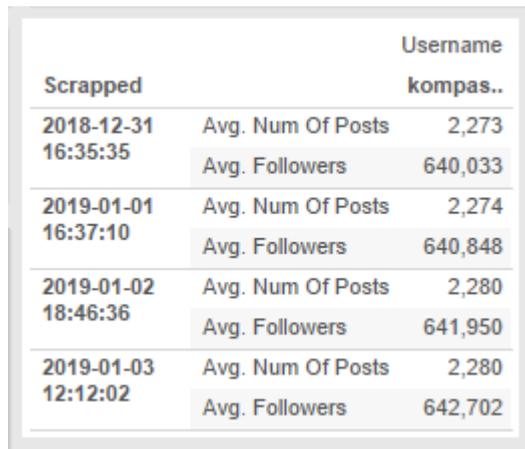
Scrapped		Username
		detikcom
2018-12-31 16:26:10	Avg. Num Of Posts	3,192
	Avg. Followers	1,135,647
2019-01-01 16:27:10	Avg. Num Of Posts	3,208
	Avg. Followers	1,137,620
2019-01-02 18:36:56	Avg. Num Of Posts	3,213
	Avg. Followers	1,140,277
2019-01-03 12:02:16	Avg. Num Of Posts	3,215
	Avg. Followers	1,141,919

Gambar 6.5 Perbandingan jumlah *followers* dan *posts* detikcom

Gambar 6.5 menjelaskan mengenai perubahan aktifitas dan kepopuleran akun detikcom dalam empat hari mulai dari tanggal 31 Desember 2018 sampai 3 Januari 2019. Detikcom pada tanggal tiga puluh satu Desember 2018 mempunyai jumlah *followers* mencapai satu juta seratus tiga puluh lima ribu enam ratus empat puluh tujuh dengan jumlah *posts* yang dibuat adalah tiga ribu Sembilan puluh dua. Pada tanggal satu Januari 2019 terdapat tiga ribu dua ratus delapan *post* dengan jumlah *followers* satu juta seratus tiga puluh tujuh ribu enam ratus dua puluh.

Pada pengambilan data tanggal dua Januari 2019 terdapat tiga ribu dua ratus tiga belas *posts* dengan jumlah *followers* satu juta seratus empat puluh ribu dua ratus tujuh puluh tujuh. Dan pada tanggal tiga mempunyai jumlah *post* tiga ribu dua ratus lima

belas dengan jumlah *followers* mencapai satu juta seratus empat puluh satu ribu Sembilan ratus Sembilan belas.



Scrapped		Username
2018-12-31 16:35:35	Avg. Num Of Posts	2,273
	Avg. Followers	640,033
2019-01-01 16:37:10	Avg. Num Of Posts	2,274
	Avg. Followers	640,848
2019-01-02 18:46:36	Avg. Num Of Posts	2,280
	Avg. Followers	641,950
2019-01-03 12:12:02	Avg. Num Of Posts	2,280
	Avg. Followers	642,702

Gambar 6.6 Perbandingan jumlah *followers* dan *posts* kompascom

Pada empat hari pengambilan data, kompascom mempunyai *followers* sebanyak enam ratus empat puluh ribu tiga puluh tiga dengan jumlah *post* sebanyak dua ribu dua ratus tujuh puluh tiga pada pengambilan data tanggal tiga puluh satu Desember 2018. Pada tanggal satu Januari 2019 jumlah *followers* bertambah menjadi enam ratus empat puluh ribu delapan ratus empat puluh delapan sementara jumlah *post* menjadi dua ribu dua ratus tujuh puluh empat.

Pada tanggal dua Januari jumlah *posts* bertambah menjadi dua ribu dua ratus delapan puluh dan jumlah *followers* bertambah menjadi enam ratus empat puluh satu ribu Sembilan ratus lima puluh. Sementara pada pengambilan data tanggal tiga Januari jumlah *posts* tidak bertambah sementara jumlah *followers* menjadi enam ratus empat puluh dua ribu tujuh ratus dua. Gambar 6.6 menunjukkan tabel pada tableau yang menunjukkan jumlah *post* dan *followers* kompascom.

Scrapped		Username
		indone..
2018-12-31 16:39:03	Avg. Num Of Posts	1,879
	Avg. Followers	61,584
2019-01-01 16:40:26	Avg. Num Of Posts	1,881
	Avg. Followers	61,839
2019-01-02 18:50:22	Avg. Num Of Posts	1,882
	Avg. Followers	62,181
2019-01-03 12:15:20	Avg. Num Of Posts	1,884
	Avg. Followers	62,452

Gambar 6.7 Perbandingan jumlah *followers* dan *posts* sindonews

Pada pengambilan data tanggal tiga puluh satu sindonews mempunyai jumlah *post* sebanyak seribu delapan ratus tujuh puluh Sembilan dengan jumlah *follower* sebanyak enam puluh satu ribu lima ratus delapan puluh empat. Pada satu Januari jumlah *post* bertambah menjadi seribu delapan ratus delapan puluh satu dan jumlah *followers*.

Pada pengambilan data tanggal dua Januari sindonews mempunyai jumlah *post* sebanyak seribu delapan ratus delapan puluh dua dengan jumlah *follower* sebanyak enam puluh dua ribu seratus delapan puluh satu. Pada tiga Januari jumlah *post* bertambah menjadi seribu delapan ratus delapan puluh empat dan jumlah *followers* menjadi enam puluh dua ribu empat ratus lima puluh dua. Gambar 6.7 menunjukkan keseluruhan data dari akun sindonews yang sudah diambil untuk penelitian kali ini.

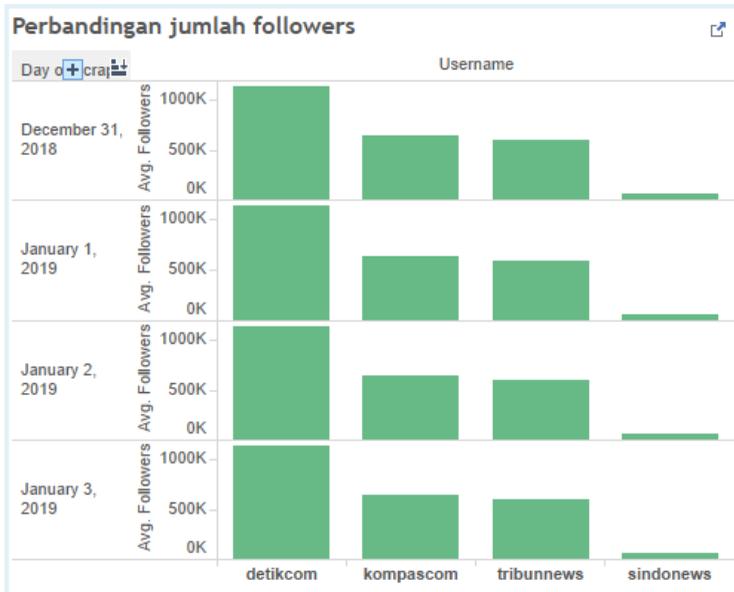
Scrapped		Username
		tribunne..
2018-12-31 17:03:21	Avg. Num Of Posts	4,995
	Avg. Followers	597,861
2019-01-01 17:06:46	Avg. Num Of Posts	5,001
	Avg. Followers	598,897
2019-01-02 19:11:38	Avg. Num Of Posts	5,009
	Avg. Followers	600,342
2019-01-03 12:39:24	Avg. Num Of Posts	5,015
	Avg. Followers	601,225

Gambar 6.8 Perbandingan jumlah *followers* dan *posts* tribunnews

Gambar 6.8 menjelaskan mengenai perubahan aktifitas dan kepopuleran akun tribunnews dalam empat hari mulai dari tanggal tiga puluh satu Desember dua ribu delapan belas sampai tiga Januari dua ribu Sembilan belas. Detikcom pada tanggal tiga puluh satu Desember dua ribu delapan belas mempunyai jumlah *followers* mencapai lima ratus Sembilan puluh tujuh ribu delapan ratus enam puluh satu dengan jumlah *posts* yang dibuat adalah empat ribu Sembilan ratus Sembilan puluh lima.

Pada tanggal satu Januari 2019 terdapat lima ribu satu *post* dengan jumlah *followers* lima ratus Sembilan puluh delapan ribu delapan ratus Sembilan puluh tujuh. Pada pengambilan data tanggal dua Januari 2019 terdapat lima ribu sembilan *posts* dengan jumlah *followers* enam ratus ribu tiga ratus empat puluh dua. Dan pada tanggal tiga mempunyai jumlah *post* lima ribu dua ratus lima belas dengan jumlah *followers* mencapai enam ratus seribu dua ratus dua puluh lima. Total *post* yang dibuat oleh tribunnews pada periode pengambilan data adalah sebanyak 20 *post* baru.

6.1.3. Grafik Perbandingan *Followers* dan *Post*

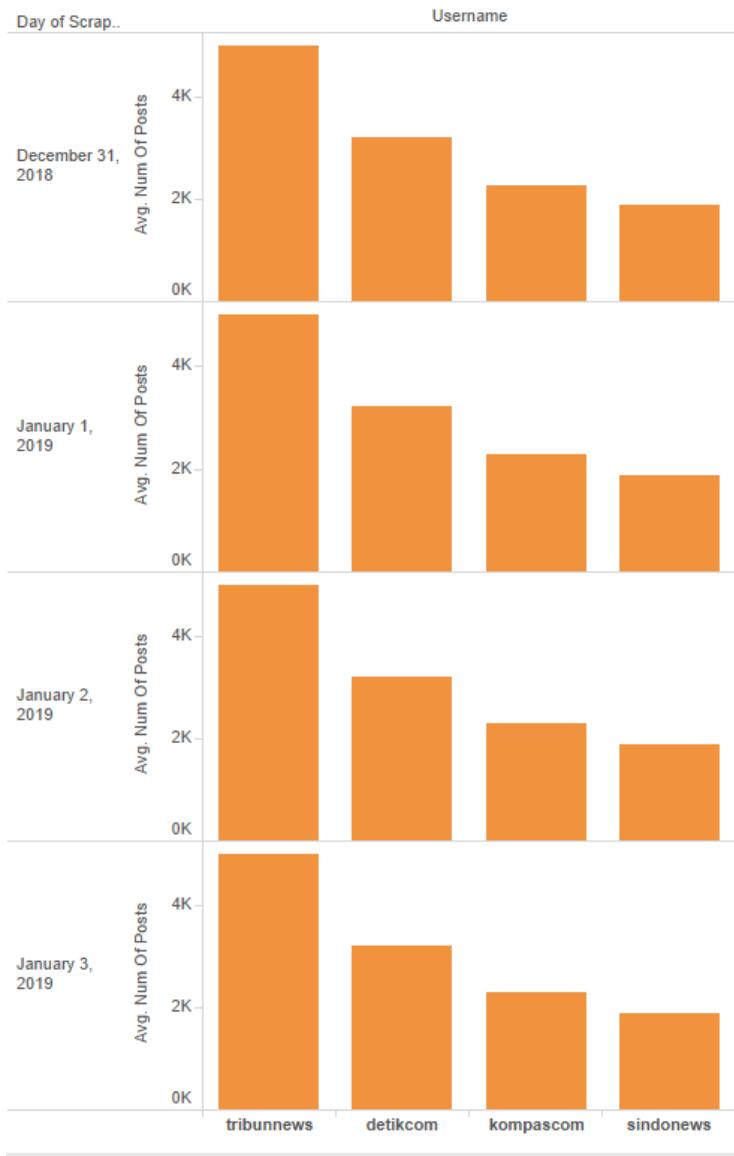


Gambar 6.9 Perbandingan jumlah *followers*

Gambar 6.9 merupakan grafik yang menunjukkan perbandingan jumlah *followers* untuk empat akun yang diambil datanya. Data diatas menunjukkan sampel yang lebar dengan detikcom teratas dan kompascom dan tribunews menjadi akun pada titik tengah, sementara sindonews memiliki basis *followers* yang rendah sekali hanya sepersepuluh dari kompascom dan tribunews. Detikcom pada 31 desember mempunyai satu juta lebih *followers* dan menempati peringkat paling atas sebagai akun dengan jumlah *followers* paling banyak.

Sementara kompascom berada di posisi kedua sebagai akun dengan *followers* terbanyak diikuti oleh tribunews yang mempunyai *followers* dengan selisih beberapa puluh ribu diantara keduanya. Sindonews hanya dapat berada di peringkat paling buncit dengan hanya mempunyai *followers* pada kisaran enam puluhan ribu. Pertumbuhan jumlah *followers* tidak ada lonjakan yang signifikan pada tiga hari berikutnya.

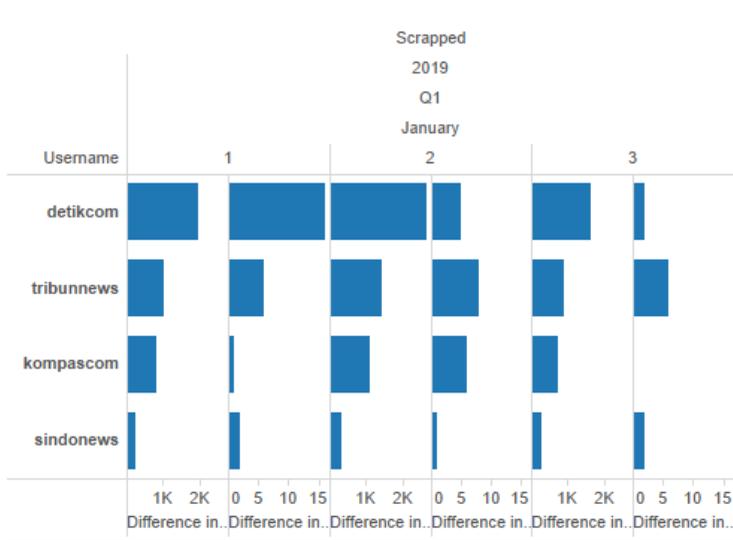
Perbandingan jumlah post



Gambar 6.10 Perbandingan jumlah post

Gambar 6.10 menunjukkan jumlah *post* yang dibuat oleh ke empat akun tersebut. Tribunnews mempunyai jumlah *post* paling banyak yang mencapai lima ribu jumlah *post* pada pengambilan data. Sementara detikcom mempunyai jumlah *post* yang mencapai lebih dari tiga ribu *post*.

Sementara kompascom mempunyai sekira dua ribu *post* dan sindonews tetap berada pada posisi paling rendah dari segi keaktifan karena hanya membuat *post* kurang dari dua ribu.



Gambar 6.11 Perbandingan perubahan *post* dan *followers*

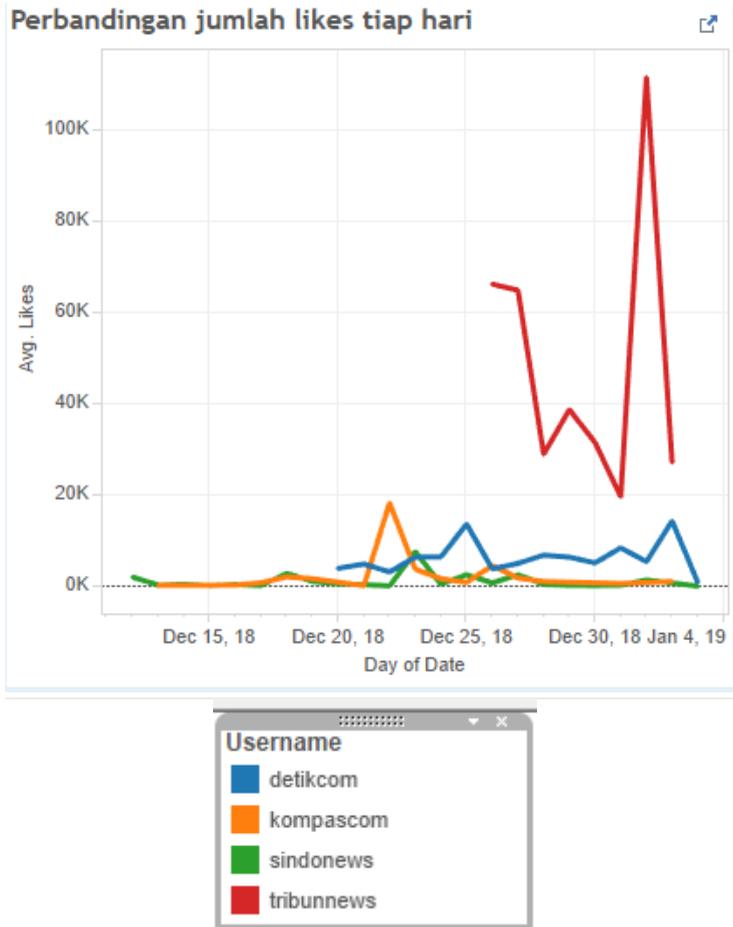
Pada gambar 6.11 menunjukkan perubahan jumlah *followers* dan *post* tiap kali melakukan pengambilan data dengan pengambilan data tanggal 31 Desember 2018 sebagai *benchmark*. Pada periode tanggal tiga puluh satu Desember 2018 sampai dengan pengambilan data pada tanggal satu Januari 2019 detikcom mempunyai *followers* baru hampir mencapai dua ribu akun *followers* baru dan membuat *post* sebanyak 16 *post* baru.

Hal ini bisa didapatkan dengan membandingkan jumlah *followers* pada tanggal satu Januari dengan jumlah *followers*

pada tanggal tiga puluh satu Desember 2018. Tribunnews mempunyai lebih dari seribu *followers* baru dan lima *post* baru antara periode tiga puluh satu Desember 2018 dan satu Januari 2019. Kompascom mempunyai sekitar seribu *followers* baru dan satu *post* baru antara periode tiga puluh satu Desember 2018 dan satu Januari 2019. Sindonews mempunyai kurang dari lima ratus *followers* baru dan dua *post* baru antara periode tiga puluh satu Desember 2018 dan satu Januari 2019.

Sementara untuk periode antara tanggal satu Januari dan dua Januari detikcom mempunyai *followers* baru kurang dari dua ribu akun *followers* baru dan membuat *post* sebanyak tiga *post* baru. Hal ini bisa didapatkan dengan membandingkan jumlah *followers* pada tanggal satu Januari dengan jumlah *followers* pada tanggal dua Januari. Tribunnews mempunyai lebih dari seribu *followers* baru dan tujuh *post* baru antara periode satu Januari 2019 dan dua Januari 2019. Kompascom mempunyai lebih dari seribu *followers* baru dan lima *post* baru antara periode satu Januari 2018 dan dua Januari 2019. Sindonews mempunyai kurang dari lima ratus *followers* baru dan satu *post* baru antara periode satu Januari 2018 dan dua Januari 2019.

Sementara untuk periode antara tanggal dua Januari dan tiga Januari detikcom mempunyai *followers* baru lebih dari dua ribu akun *followers* baru dan membuat *post* sebanyak dua *post* baru. Hal ini bisa didapatkan dengan membandingkan jumlah *followers* pada tanggal dua Januari dengan jumlah *followers* pada tanggal tiga Januari. Tribunnews mempunyai lebih dari seribu *followers* baru dan enam *post* baru antara periode dua Januari 2019 dan tiga Januari 2019. Kompascom mempunyai kurang dari seribu *followers* baru dan tidak membuat *post* baru antara periode dua Januari 2019 dan tiga Januari 2019. Sindonews mempunyai kurang dari lima ratus *followers* baru dan membuat dua *post* baru antara periode dua Januari 2019 dan tiga Januari 2019.



Gambar 6.12 Perbandingan jumlah *likes* tiap akun

Gambar 6.12 merupakan rata rata jumlah *likes* tiap hari yang diterima oleh akun yang diterangkan dalam gambar 6.13. Grafik ini menunjukkan ketertarikan khalayak terhadap ke empat akun tersebut. Tribunnews mempunyai tingkat interaksi yang tinggi dengan memiliki jumlah *likes* tiap harinya melebihi ke tiga akun lain. Tribunnews mempunyai jumlah *likes* yang lebih besar tiap harinya tetapi mereka ketertarikan terhadap *post* yang dibuat oleh tribunnews naik turun secara signifikan. Detikcom

mempunyai jumlah *likes* yang lebih stabil dan cukup tinggi dengan tiap harinya mendapat sepuluh ribu *likes* jika dibandingkan dengan ketiga akun yang lain.

6.2. *Followers Growth Analytics*

Followers growth analytics pada penelitian kali ini adalah untuk melihat bagaimana pertumbuhan jumlah *followers* pada periode tertentu dibandingkan dengan jumlah *posts*. Komponen yang dikembangkan dalam halaman dashboard antara lain:

- Perbandingan jumlah *followers*
- Perbandingan jumlah *posts*
- Perbandingan jumlah *followers* dan *post* pada periode tertentu.
- Perbandingan jumlah *likes* pada *posts* pada periode tertentu.
- Filter *username*

Gambar 6.13 merupakan dashboard yang merupakan kumpulan dari grafik yang ada pada bagian sebelumnya.

Dashboard ini membicarakan mengenai perkembangan jumlah *followers* dari detikcom, tribunnews, kompascom dan sindonews yang bisa dilihat pada grafik kiri atas yang mempunyai warna hijau.

Perkembangan jumlah *post* dari detikcom, tribunnews, kompascom dan sindonews yang bisa dilihat pada grafik kanan atas yang berwarna oranye.

Sementara grafik perbandingan pertumbuhan *followers* dan aktifitas akun detikcom, tribunnews, kompascom dan sindonews bisa dilihat pada pojok kiri dashboard berwarna biru. Sementara grafik pojok kanan pada dashboard berisi mengenai tingkat interaksi *post* tersebut dengan akun lain yang ditunjukkan dengan jumlah *likes* yang diterima.



Gambar 6.13 Dashboard *followers growth analytics*.

6.3. Analisis Pertumbuhan *Followers* dan *Post*

Dari tabel *username* dan tanggal *scraped* yang mempunyai jumlah *post* dan jumlah *followers* pada subbab sebelumnya ada beberapa analisis yang dapat diberikan.

Analisis dari akun terpopuler dan dengan aktifitas terbanyak sepanjang umur akun tersebut adalah sebagai berikut.

1. Detikcom merupakan akun berita dengan jumlah *followers* terbanyak pada pengambilan data kali ini
2. Tribunnews merupakan akun berita dengan jumlah *posts* terbanyak pada pengambilan data kali ini.

Sementara jika di analisis lebih lanjut aktifitas yang terjadi pada akun tersebut dan pertumbuhan jumlah *followers* berikut analisisnya.

	Scrapped	
	2018	2019
	Q4	Q1
	December	January
Username₂₊	31	3
detikcom	0.00	23.00
tribunnews	0.00	20.00
kompascom	0.00	7.00
sindonews	0.00	5.00

Gambar 6.14 Jumlah *post* yang dibuat

Gambar 6.14 menunjukkan perubahan jumlah *post* yang terjadi pada periode tiga puluh satu Desember 2018 sampai tiga Januari 2019. Detikcom merupakan akun teraktif dengan membuat dua puluh tiga *post* pada periode pengambilan data. Sementara pada tribunnews membuat dua puluh *post*, kompascom dengan tujuh *post* baru dan sindonews menjadi akun dengan aktifitas terendah.

Hal ini tidak beda jauh dengan jumlah total *posts* yang dimiliki oleh ke empat akun tersebut dengan perbedaan detikcom lebih aktif pada periode ini dengan selisih tiga *post*. Akan tetapi akun dengan jumlah *post* tertinggi masih dipegang oleh tribunnews dengan 5015 *posts* sangat jauh sekali dengan detikcom yang hanya memiliki 3215 *post* baru.

	Scrapped	
	2018	2019
	Q4	Q1
	December	January
Username	31	3
detikcom	0	6,272
tribunnews	0	3,364
kompascom	0	2,669
sindonews	0	868

Gambar 6.15 Jumlah *followers* yang didapat

Gambar 6.15 menunjukkan perbandingan jumlah *followers* pada tiga puluh satu Desember 2018 sampai dengan tiga Januari 2019. Detikcom merupakan akun dengan pertumbuhan terbesar pada periode pengambilan data kali ini dengan pertumbuhan *followers* sebesar 6.272 diikuti oleh tribunnews dengan 3.364 pada peringkat kedua dan kompascom dengan 2.669 *followers* baru dan sindonews berada pada posisi buncit dengan 868 *followers* baru.

Tetapi jika dibandingkan dengan prosentase pertambahan *followers* baru dengan basis *followers* yang sudah dimiliki sindonews memiliki pertumbuhan yang paling tinggi dengan nilai di atas satu prosen, di ikuti oleh tribunnews dengan pertumbuhan nol koma lima puluh enam persen pada peringkat kedua. Detikcom mempunyai prosentase pertumbuhan sebesar nol koma lima puluh lima. Sementara kompascom mempunyai prosentase pertumbuhan terendah dengan hanya nol koma empat puluh dua prosen. Untuk selengkapnya dapat dilihat pada gambar 6.16.

Username	Scrapped	
	2018	2019
	Q4 December	Q1 January
	31	3
sindonews	0.000%	1.409%
tribunnews	0.000%	0.563%
detikcom	0.000%	0.552%
kompascom	0.000%	0.417%

Gambar 6.16 Prosentase pertumbuhan *followers*

Username		Scrapped	
		2018	2019
		Q4 December	Q1 January
		31	3
detikcom	% Difference in Avg. Follower..	0.00%	0.55%
	Difference in Avg. Followers fr..	0	6,272
tribunnews	% Difference in Avg. Follower..	0.00%	0.56%
	Difference in Avg. Followers fr..	0	3,364
kompascom	% Difference in Avg. Follower..	0.00%	0.42%
	Difference in Avg. Followers fr..	0	2,669
sindonews	% Difference in Avg. Follower..	0.00%	1.41%
	Difference in Avg. Followers fr..	0	868

Gambar 6.17 Prosentase perubahan *followers* dan jumlah perubahan *followers*

Gambar 6.17 menunjukkan prosentase perubahan *followers* dengan jumlah perubahan *followers*. Memang secara angka kasar detikcom mempunyai jumlah pertambahan paling tinggi tetapi secara prosentase detikcom hanya bertumbuh 0,55 prosen. Tidak jauh beda dengan tribunnews yang punya

pertumbuhan jumlah *followers* sebesar 3.364 tapi prosentase pertumbuhannya sebesar 0,56 persen. Yang paling mengejutkan adalah prosentase pertumbuhan sindonews yang mencapai 1,41 persen hanya dengan 868 *followers* baru.

Sementara kompascom hanya dapat mencapai prosentase pertumbuhan sebesar 0,42 persen dengan jumlah *followers* baru sebesar 2669. Hal ini dapat terjadi karena jumlah *followers* data awal yang jauh berbeda detikcom mempunyai *followers* sebesar satu juta lebih jadi ketika terjadi penambahan enam ribuan *followers* tidak terlalu besar. Beda halnya dengan sindonews yang hanya mempunyai jumlah *followers* awal sebesar enam puluhan ribu. Hal ini menyebabkan prosentase pertumbuhan yang tinggi.

		Scrapped	
		2018 Q4 December	2019 Q1 January
Username		31	3
detikcom	Difference in Avg. Num Of Po..	0	23
	Difference in Avg. Followers fr..	0	6,272
tribunnews	Difference in Avg. Num Of Po..	0	20
	Difference in Avg. Followers fr..	0	3,364
kompascom	Difference in Avg. Num Of Po..	0	7
	Difference in Avg. Followers fr..	0	2,669
sindonews	Difference in Avg. Num Of Po..	0	5
	Difference in Avg. Followers fr..	0	868

Gambar 6.18 Perbandingan jumlah *followers* dan *post*

Gambar 6.18 menunjukkan perbandingan jumlah *followers* dan *post* yang dibuat. Detikcom merupakan akun dengan aktifitas terbanyak dengan melakukan 23 *posts*. Karena mempunyai aktifitas yang besar detikcom juga punya pertambahan jumlah *followers* terbanyak dengan 6272 *followers*.

Sementara kompascom walaupun Cuma membuat 7 *posts* pada periode ini atau akun dengan aktifitas terendah nomor 2 setelah

sindonews mempunyai *followers* baru sebanyak 2669. Tribunnews membuat 20 *post* baru pada periode ini dengan penambahan 3364 *followers* baru. Sindonews yang mempunyai 868 *followers* baru hanya membuat lima *post* baru pada periode pengambilan data.

Analisis lain yang mampu ditarik dari jumlah *post* dan penambahan jumlah *followers* adalah sebagai berikut

1. Jika dibandingkan dengan jumlah *post*-nya detikcom hanya mendapat 273 *followers* baru tiap *post* yang dibuat.
2. Tiap kali kompascom membuat *post* baru mampu menarik perhatian 381 *followers* baru.
3. Tribunnews lebih aktif sehingga hanya mampu membuat 168 *followers* baru tiap *post*-nya.
4. Karena *reach* yang dimiliki oleh sindonews lebih rendah daripada akun lain membuat pertumbuhan jumlah *followers*-nya pun terendah dibandingkan akun lain. Hanya 868 *followers* baru pada periode pengambilan data.
5. Tapi dari jumlah *followers* yang didapat jika dibandingkan dengan jumlah *post*-nya masih lebih baik daripada tribunnews dengan 174 *followers* baru dibandingkan 168 *followers* baru.
6. Beda halnya dengan kompascom yang mampu mendapat 381 *followers* tiap *post* yang dibuatnya. Hal ini menunjukkan *post* kompascom lebih berharga daripada tribunnews.
7. Tribunnews mempunyai rasio *followers* per *posts* terjelek dengan 168 *followers* baru tiap kali membuat *posts*.

Username		Scrapped	
		2018	2019
		Q4 December	Q1 January
		31	3
detikcom	Difference in Avg. Num Of Po..	0	23
	Difference in Avg. Followers fr..	0	6,272
	Avg. Followers	1,135,647	1,141,919
tribunnews	Difference in Avg. Num Of Po..	0	20
	Difference in Avg. Followers fr..	0	3,364
	Avg. Followers	597,861	601,225
kompascom	Difference in Avg. Num Of Po..	0	7
	Difference in Avg. Followers fr..	0	2,669
	Avg. Followers	640,033	642,702
sindonews	Difference in Avg. Num Of Po..	0	5
	Difference in Avg. Followers fr..	0	868
	Avg. Followers	61,584	62,452

Gambar 6.19 Pertumbuhan akun dengan basis *followers*

Pada periode ini walaupun *reach* yang dimiliki oleh kompascom dan tribunnews hampir sama yaitu pada 600an ribu pengikut, pertumbuhan pengikut tribunnews lebih tinggi, yaitu 3364 *followers* baru dibandingkan 2669 *followers* baru.

Sementara jika di analisis berdasarkan tingkat *reach* atau *pool followers*. *Reach* disini membicarakan mengenai jumlah *followers* yang sudah dimiliki oleh akun tersebut. Hal ini jelas menguntungkan bagi akun dengan *followers* yang lebih tinggi karena algoritma yang digunakan oleh instagram yang menampilkan *post* baru ke lebih banyak lini masa.

Yang paling jelas adalah pertumbuhan akun salah satunya adalah karena dipengaruhi *reach* atau *pool followers*. Dengan *followers* mencapai 1 juta detikcom mempunyai basis yang kuat

sehingga pertumbuhannya mencapai 6 ribuan jumlah *followers* dalam periode pengambilan data.

Sementara kompascom dan tribunnwes mempunyai *reach* hampir mirip sehingga perubahannya pun mirip satu sama lain. Tetapi karena *post* yang dibuat lebih banyak tribunnews pun mendapat *followers* yang lebih banyak daripada kompascom sebesar 3364 dibandingkan 2669 yang diperoleh kompascom pada periode yang sama. Karena *reach*-nya lebih kecil, sindonews pun mengalami pertumbuhan *followers* paling sedikit pada periode pengambilan data ini yaitu hanya sebesar 868 *followers*.

(Halaman ini sengaja dikosongkan)

BAB VII

KESIMPULAN DAN SARAN

Bab ini berisikan kesimpulan dari hasil penelitian dan juga saran perbaikan untuk penelitian kedepannya.

7.1. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan terdapat beberapa kesimpulan yang dapat diambil

1. Perlu dilakukan *crawling* menggunakan *brute force* untuk mengambil data dari *instagram* versi web
2. Data data penting yang dapat diambil dari *crawling* intstagram versi web antara lain *username*, jumlah *post*, jumlah *followers*, jumlah *following*, tanggal pengambilan data, tanggal melakukan *posting* dan jumlah *likes* serta *user* pada tiap *post*
3. Dalam pengembangan dashboard terdapat data yang belum dimasukan dalam dashboard seperti komentar dan isi *posts*.
4. Dashboard yang dibuat terlalu umum dan sulit dibaca.
5. Detikcom merupakan akun dengan jumlah *followers* tertinggi.
6. Tribunnews merupakan akun dengan jumlah *likes* tertinggi.
7. Akun dengan pertumbuhan *followers* terbanyak adalah detikcom.
8. Selain akun dengan pertumbuhan *followers* terbanyak *detikcom* juga membuat *post* terbanyak.
9. Kompascom merupakan akun dengan *posts* paling berharga, karena sekali *posts* akun kompascom dapat membuat 381 *followers* baru.
10. Jumlah data yang diambil terlalu sedikit untuk membuat kesimpulan apakah jumlah *post* yang dibuat pada periode tertentu mempengaruhi pertumbuhan sebuah akun.

7.2. Saran

Saran yang dapat dipertimbangkan untuk pengembangan lebih lanjut penelitian ini adalah sebagai berikut:

1. Penelitian selanjutnya dapat dilakukan untuk melihat sentimen pada komentar yang diberikan oleh *user* pada *post* sebuah akun.
2. Visualisasi data dapat lebih dikembangkan untuk melihat apakah ada korelasi antara waktu *posting* dengan interaksi pada *post* sebuah akun instagram.

DAFTAR PUSTAKA

- [1] O. Kwon and Y. Wen, "An empirical study of the factors affecting social network service use," *Computers in Human Behaviour*, vol. XXVI, pp. 254-263, 2010.
- [2] Statista, "Number of monthly active Instagram *users* from January 2013 to September 2017 (in millions)," Statista, 11 Februari 2018. [Online]. Available: <https://www.statista.com/statistics/253577/number-of-monthly-active-instagram-users/>. [Accessed 19 Februari 2018].
- [3] S. Aslam, "Omnicoagency Instagram Statistics," Omnicore group, 1 January 2018. [Online]. Available: <https://www.omnicoreagency.com/instagram-statistics/>. [Accessed 23 3 2018].
- [4] B. D. Saputri, "Kompasiana," *kompas*, 2 October 2017. [Online]. Available: <https://www.kompasiana.com/bencha/59d0da972ba8d16a6211c7f2/peran-media-sosial-instagram-dalam-penyebaran-berita?page=all>. [Accessed 9 Januari 2019].
- [5] V. Mittal, A. Kaul, S. S. Gupta and A. Arora, "Multivariate Features Based Instagram *Post* Analysis to Enrich *User* Experience," *Procedia Computer Science*, no. 122, pp. 138-145, 2017.
- [6] R. Liu and A. Suh, "Self-Branding on Social Media: An Analysis of Style Bloggers on Instagram," *Procedia Computer Science*, no. 124, pp. 12-20, 2017.
- [7] E. Djafarova and C. Rushworth, "Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female *users*," *Computers in Human Behavior*, vol. 68, pp. 1-7, 2017.

- [8] Y. Hu, L. Manikonda and S. Kambhampati, "What we Instagram: a first analysis of instagram photo content and *user* types," in *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*, Ann Arbor, 2014.
- [9] B. Jensen, "Instagram as Cultural Heritage," in *Digital Heritage International Congress* , Marseille, France, 2013.
- [10] C. R. King and R. K. Freeman, "FLEXIBLE DATABASE SCHEMA". United States of America Patent US 9,031.901 B1 , 12 May 2015.
- [11] V. Rainardi, *Building a data warehouse: with examples in SQL server*, John Wiley and sons, 2011.
- [12] R. Elmasri and S. Navathe, *Fundamentals of Database Systems*, Addison Wesley, 2010.
- [13] Python, "Python," Python Software Foundation, 4 March 2018. [Online]. Available: <https://docs.python.org/3/faq/general.html>. [Accessed 4 March 2018].
- [14] C. Hanretty, "*Scraping* the web for arts and humanities," University of East Anglia, 2013.
- [15] MySQL, "MySQL About," Oracle, 2018. [Online]. Available: <https://www.mysql.com/about/>. [Accessed 12 3 2018].
- [16] J. Martin, *Rapid Application Development*, Macmillan, 1991.
- [17] P. Beynon-Davies, H. Mackay, C. Carne and D. Tudhope, "Rapid application development (RAD): An empirical review," *European Journal of Information Systems*, vol. 8, no. 3, 1999.
- [18] D. Hom, R. Perez and L. Williams, "Which chart or graph is right for you," 2016, Tableau Software.

- [19] D. Crockford, "Introducing JSON," Json.org, [Online]. Available: <https://www.json.org/>. [Accessed 24 1 2019].
- [20] M. Bean, Laravel 5 Essentials, Packt Publishing, 2015.

(halaman ini sengaja dikosongkan)

LAMPIRAN A SOURCE CODE

Modul 1 *Crawler* instagram versi web

<i>Script</i>	Keterangan	Lokasi
Crawl_profile.py	Melakukan pengambilan data	\instagram-profilecrawl-master\crawl_profile.py
Account.py	Melakukan login pada instagram versi web	\instagram-profilecrawl-master\util\account.py
Chromedriver.py	Melakukan otomasi browser chrome	instagram-profilecrawl-master\util\chromedriver.py
Cli_helper.py	Menyiapkan <i>username</i> yang akan dicari	instagram-profilecrawl-master\util\cli_helper.py
Datasaver.py	Menyimpan dengan format Json	instagram-profilecrawl-master\util\datasaver.py
Extractor.py	Mengambil elemen dari halaman instagram	instagram-profilecrawl-master\util\extractor.py
Settings.py	Pengaturan dan batasan <i>crawling</i>	instagram-profilecrawl-master\util\settings.py

Modul 2 import data

<i>Script</i>	<i>Keterangan</i>	<i>Lokasi</i>
<i>Import_username.php</i>	Memasukan data ke tabel <i>user_account</i>	instagram-profilecrawl-master\db\Import_username.php
<i>Import_post.php</i>	Memasukan data ke tabel <i>posts</i>	instagram-profilecrawl-master\db\import_post.php
<i>Import_comments.php</i>	Memasukan data ke tabel <i>comments</i>	instagram-profilecrawl-master\db\import_comments.php

BIODATA PENULIS



Penulis lahir di Boyolali pada tanggal 5 September 1994. Penulis merupakan anak kedua dari dua bersaudara. Penulis telah menempuh pendidikan formal di sekolah negeri mulai dari SD Jenengan 3 hingga tahun 2006, SMPN 1 Sawit dan SMAN 1 Boyolali. Setelah lulus, penulis melanjutkan ke jenjang perguruan tinggi negeri di Surabaya, yakni Jurusan Sistem Informasi Institut Teknologi Sepuluh

Nopember Surabaya

Untuk mendapatkan gelar Sarjana Komputer (S.Kom), penulis mengambil laboratorium bidang minat Sistem Enterprise (SE) dengan topik tugas akhir *crawling* data instagram versi web. Untuk kepentingan penelitian penulis juga dapat dihubungi melalui e-mail: afpras@gmail.com.