

**TEKNIK WEB SCRAPING UNTUK ANALISIS  
HUBUNGAN POPULARITAS SEIYUU DAN ANIME  
BERDASARKAN DATA MYANIMELIST**

**SKRIPSI SARJANA**

Oleh:

Raihan Tsabita Sabil

207064516082



**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI KOMUNIKASI DAN  
INFORMATIKA  
UNIVERSITAS NASIONAL  
2024**

**TEKNIK WEB SCRAPING UNTUK ANALISIS  
HUBUNGAN POPULARITAS SEIYUU DAN ANIME  
BERDASARKAN DATA MYANIMELIST**

**SKRIPSI SARJANA**

Karya Ilmiah salah satu syarat untuk memperoleh gelar sarjana  
Informatika dari Fakultas Teknologi Komunikasi Dan Informatika

Oleh:

Raihan Tsabita Sabil

207064516082



**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI KOMUNIKASI DAN  
INFORMATIKA  
UNIVERSITAS NASIONAL**

**2024**

HALAMAN PENGESAHAN

TUGAS AKHIR

TEKNIK WEB SCRAPING UNTUK ANALISIS HUBUNGAN  
POPULARITAS SEIYUU DAN ANIME BERDASARKAN DATA  
MYANIMELIST



(Ratih Titi Komala Sari, S.T., M.M., MMSI)

## HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Bilamana di kemudian hari ditemukan bahwa karya tulis ini menyalahi peraturan yang ada berkaitan etika dan kaidah penulisan karya ilmiah yang berlaku, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku

Yang menyatakan,

Nama : Raihan Tsabita Sabil

NIM : 207064516082

Tanda Tangan :



Tanggal : Kamis, 22 Agustus 2024

Mengetahui

Pembimbing I : Ratih Titi Komala Sari, ST. MM., MMSI (

A handwritten signature in black ink, likely belonging to Ratih Titi Komala Sari, is written inside a pair of parentheses.

**HALAMAN PENGESAHAN**  
**TUGAS SARJANA**  
**TEKNIK WEB SCRAPING UNTUK ANALISIS HUBUNGAN**  
**POPULARITAS SEIYUU DAN ANIME BERDASARKAN DATA**  
**MYANIMELIST**

Oleh:

Raihan Tsabita Sabil

207064516082

Program Studi Informatika  
Fakultas Teknologi Komunikasi dan Informatika

Disetujui pada tanggal: 22 Agustus 2024

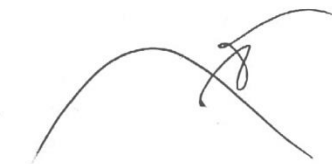
Pembimbing 1



Ratih Titi Komala Sari, ST. MM., MMSI  
NIDN. 0301038302

Penguji 1

Penguji 2



Novi Dian Nathasia, S.Kom., MMSI  
NIDN. 0714127701



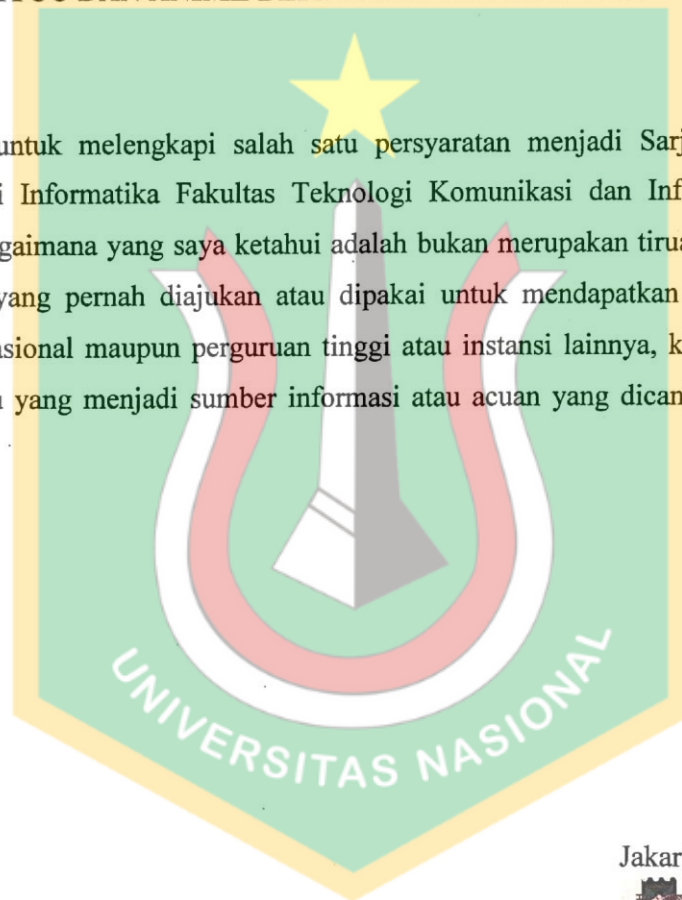
Dr. Benrahman, S.Kom., M.MSI  
NIDN. 0318096504

## PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

### TEKNIK WEB SCRAPING UNTUK ANALISIS HUBUNGAN POPULARITAS SEIYUU DAN ANIME BERDASARKAN DATA MYANIMELIST

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Informatika Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.



Jakarta, 27 Agustus 2024



Raihan Tsabita Sabil

207064516082

## KATA PENGANTAR

Puji dan syukur diucapkan kehadirat kepada Allah SWT yang telah memberikan rahmat dan karunia-Nya yang telah membantu penulis menyelesaikan skripsi ini sebagai salah satu syarat kelulusan dari Program Studi Sarjana Informatika di Fakultas Teknologi Komunikasi dan Informatika.

Dalam penyusunan skripsi ini, tentu tak lepas dari pengarahan dan bimbingan dari berbagai pihak. Maka penulis ucapkan rasa hormat dan terima kasih kepada semua pihak yang telah membantu. Adapun penulis juga mengucapkan banyak terimakasih kepada :

1. Kedua orang tua saya yaitu Bapak Yatimin dan Ibu Nocih Amelia yang sudah menyemangati ketika sedang suntuk maupun letih dan juga yang telah mendoakan hingga penulisan skripsi ini selesai
2. Ibu Ratih Titi Komala Sari, S.T., M.M., MMSI selaku Ketua Program Studi Teknik Informatika dan juga selaku Dosen Pembimbing Skripsi
3. Seluruh dosen pengajar di Program Studi Informatika FTKI
4. Fairuz Anwidati, Asha Maharani nopus, Ridha Afina, Radityasari, Muhammad Faiz Hafizh sebagai sahabat sejak SMA yang telah memberi banyak dukungan dan semangat juga
5. Galang dan Hafiz sebagai kawan sehoobi dan kawan kuliah.

Kami menyadari bahwa skripsi yang kami buat masih memiliki ruang untuk peningkatan, baik dari segi penyusunan, bahasa, maupun isi tulisan. Oleh karena itu, kami sangat menghargai apresiasi dan saran konstruktif dari semua pembaca. Dengan demikian, kami dapat terus berkembang dan meningkatkan kemampuan penulisan kami untuk masa yang akan datang.

Depok, 22 Agustus 2024



Raihan Tsabita Sabil



# HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI

## TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Program Studi Teknik Informatika, Fakultas Teknologi Komunikasi dan Informatika, saya yang bertanda tangan dibawah ini:

Nama : Raihan Tsabita Sabil

NIM : 207064516082

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Fakultas Teknologi Komunikasi dan Informatika, Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

### TEKNIK WEB SCRAPING UNTUK ANALISIS HUBUNGAN POPULARITAS SEIYUU DAN ANIME BERDASARKAN DATA MYANIMELIST

Beserta perangkat yang ada (jika diperlukan). Dengan hak ini Fakultas Teknologi Komunikasi dan Informatika berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pengkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilih Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok

Pada tanggal : 22 Agustus 2024

Yang menyatakan



(Raihan Tsabita Sabil)



**LEMBAR PERSETUJUAN REVIEW AKHIR**

Tugas Akhir dengan judul :

**TEKNIK WEB SCRAPING UNTUK ANALISIS  
HUBUNGAN POPULARITAS SEIYUU DAN  
ANIME BERDASARKAN DATA MYANIMELIST**


Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Informatika, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Genap 2023-2024 pada tanggal 21 Agustus Tahun 2024

**Dosen Pembimbing 1**



Ratih Titi Komala Sari, S.T., M.M., MMSI  
NID. 0103150850

**Ketua Program Studi**



Ratih Titi Komala Sari, S.T., M.M., MMSI  
NID. 0103150850

**LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI**

Nama : Raihan Tsabita Sabil  
NPM : 207064516082  
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika  
Program Studi : Informatika  
Tanggal Sidang : 21 Agustus 2024

JUDUL DALAM BAHASA INDONESIA :

**TEKNIK WEB SCRAPING UNTUK ANALISIS HUBUNGAN POPULARITAS SEIYUU DAN ANIME BERDASARKAN DATA MYANIMELIST**

JUDUL DALAM BAHASA INGGRIS :

**WEB SCRAPING TECHNIQUE FOR SEIYUU AND ANIME POPULARITY RELATIONSHIP ANALYSIS BASED ON MYANIMELIST DATA**

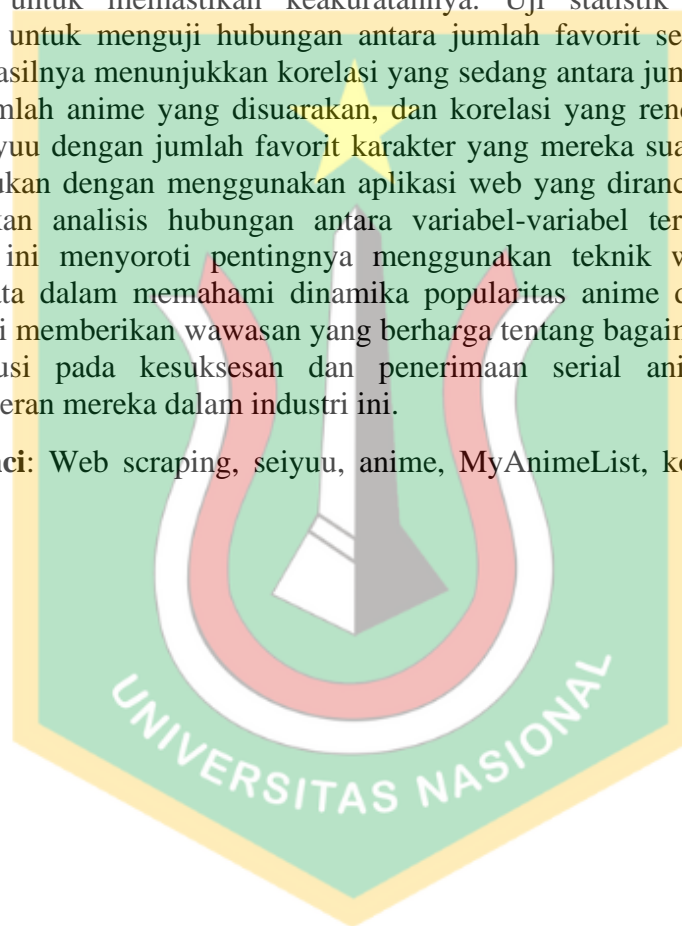
**TANDA TANGAN DAN TANGGAL**

<b>Pembimbing 1</b>	<b>Ka. Prodi</b>	<b>Mahasiswa</b>
TGL : 27 Agustus 2024	TGL : 27 Agustus 2024	TGL : 27 Agustus 2024
		

## ABSTRAK

Penelitian ini bertujuan untuk menganalisis hubungan antara popularitas seiyuu (pengisi suara) dengan anime yang mereka suarakan dengan menggunakan teknik web scraping pada data dari MyAnimeList.net. Web scraping adalah metode yang digunakan untuk mengumpulkan data secara otomatis dari halaman web. Data yang dikumpulkan meliputi nama seiyuu, jumlah favorit, nama karakter anime, jumlah karakter favorit, peringkat anime, jumlah anggota anime, dan peringkat popularitas anime. Setelah pengumpulan data, pembersihan dan pengolahan data dilakukan untuk memastikan keakuratannya. Uji statistik korelasi Pearson digunakan untuk menguji hubungan antara jumlah favorit seiyuu dan variabel lainnya. Hasilnya menunjukkan korelasi yang sedang antara jumlah favorit seiyuu dengan jumlah anime yang disuarakan, dan korelasi yang rendah antara jumlah favorit seiyuu dengan jumlah favorit karakter yang mereka suarakan. Visualisasi data dilakukan dengan menggunakan aplikasi web yang dirancang khusus untuk memudahkan analisis hubungan antara variabel-variabel tersebut. Selain itu, penelitian ini menyoroti pentingnya menggunakan teknik web scraping dan analisis data dalam memahami dinamika popularitas anime dan pengisi suara. Temuan ini memberikan wawasan yang berharga tentang bagaimana pengisi suara berkontribusi pada kesuksesan dan penerimaan serial anime, menekankan relevansi peran mereka dalam industri ini.

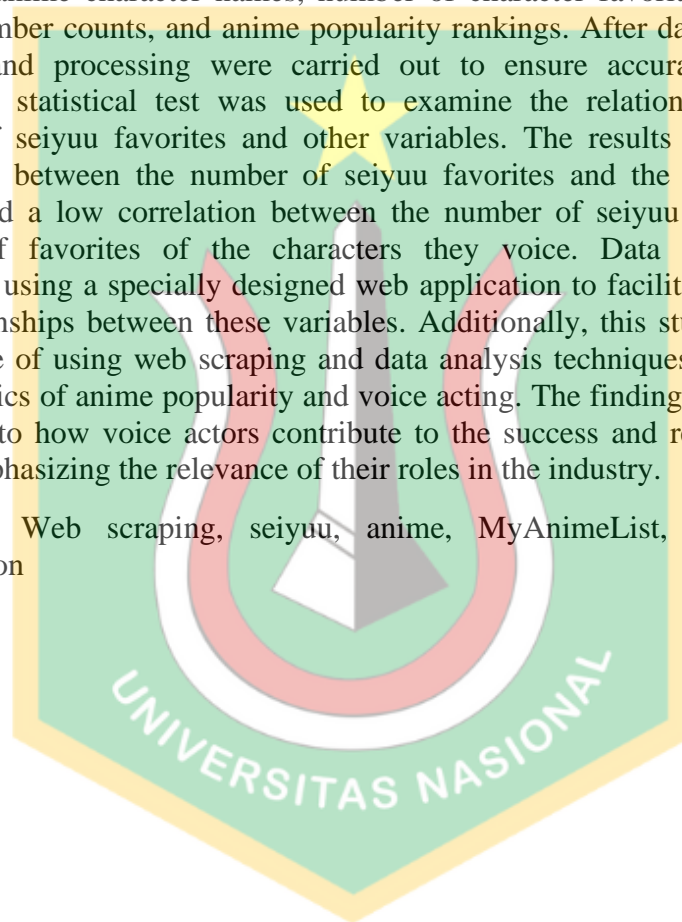
**Kata Kunci:** Web scraping, seiyuu, anime, MyAnimeList, korelasi, visualisasi data



# **Web Scraping Technique For Seiyuu And Anime Popularity Correlation Analysis Based On Myanimelist Data**

This study aims to analyze the relationship between the popularity of seiyuu (voice actors) and the anime they voice using web scraping techniques on data from MyAnimeList.net. Web scraping is a method used to automatically collect data from web pages. The collected data includes seiyuu names, number of favorites, anime character names, number of character favorites, anime ratings, anime member counts, and anime popularity rankings. After data collection, data cleaning and processing were carried out to ensure accuracy. The Pearson correlation statistical test was used to examine the relationship between the number of seiyuu favorites and other variables. The results show a moderate correlation between the number of seiyuu favorites and the number of anime voiced, and a low correlation between the number of seiyuu favorites and the number of favorites of the characters they voice. Data visualization was performed using a specially designed web application to facilitate the analysis of the relationships between these variables. Additionally, this study highlights the importance of using web scraping and data analysis techniques in understanding the dynamics of anime popularity and voice acting. The findings provide valuable insights into how voice actors contribute to the success and reception of anime series, emphasizing the relevance of their roles in the industry.

**Keyword:** Web scraping, seiyuu, anime, MyAnimeList, correlation, data visualization



## DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN TUGAS AKHIR.....	iii
HALAMAN PERNYATAAN ORISINALITAS.....	iv
HALAMAN PENGESAHAN TUGAS SARJANA .....	v
PERNYATAAN KEASLIAN TUGAS AKHIR.....	vi
KATA PENGANTAR .....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI.....	viii
LEMBAR PERSETUJUAN REVIEW AKHIR .....	ix
LEMBAR PERSETUJUAN JUDUL.....	x
ABSTRAK.....	xi
DAFTAR ISI.....	xiii
DAFTAR TABEL.....	xv
DAFTAR GAMBAR .....	xvi
DAFTAR KODE.....	xviii
BAB IPENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah.....	3
1.3 Tujuan .....	3
1.4 Batasan Masalah.....	3
1.5 Kontribusi.....	3
BAB II TINJAUAN PUSTAKAN.....	4
2.1 Web Scraping .....	4
2.2 Myanimelist.net.....	4
2.3 Anime.....	4
2.4 Seiyuu.....	5
2.5 Python .....	5
2.6 Beautifulsoup4 .....	5
2.7 HTML .....	6
2.8 Pandas .....	6
2.9 Google Colaboratory .....	6

2.10	Analisis Korelasi .....	7
2.11	Studi Literatur .....	7
2.12	Rencana Penelitian .....	15
<b>BAB III METODOLOGI PENELITIAN.....</b>		<b>16</b>
3.1	Lokasi Penelitian .....	16
3.2	Fokus Penelitian .....	16
3.3	Sumber Data.....	16
3.4	Teknik Pengumpulan Data.....	16
3.5	Tahapan Penelitian .....	17
3.5.1	Analisis kebutuhan .....	17
3.5.2	Alur Penelitian .....	17
3.5.3	Data Extracting (Pengumpulan Data) .....	23
3.5.4	Data Cleansing (Pembersihan Data) .....	35
3.5.5	Data Processing (Pemrosesan Data).....	38
3.5.6	Statistic test (Uji Statistik) .....	41
3.5.7	Data Visualization (Visualisasi Data).....	43
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>55</b>
4.1	Hasil Web Scraping.....	55
4.2	Hasil Uji Statistik.....	56
4.2.1	Hasil Uji Statistik Korelasi .....	56
4.3	Hasil Rancangan Web Aplikasi Data Visualisasi .....	59
4.4	Hasil Data Visualisasi .....	60
<b>BAB V KESIMPULAN DAN SARAN.....</b>		<b>63</b>
5.1	Kesimpulan .....	63
5.2	Saran.....	64
<b>DAFTAR PUSTAKA .....</b>		<b>65</b>
<b>LAMPIRAN.....</b>		<b>69</b>

## DAFTAR TABEL

Tabel 2.1 Daftar studi literatur .....	8
Tabel 2.2 Rencana penelitian .....	16
Tabel 4.1 Rincian jumlah data setelah scraping dan pembersihan data.....	55
Tabel 4.2 Hasil uji korelasi .....	57





## DAFTAR GAMBAR

Gambar 3.1 Flowchart alur penelitian web scraping dan analisis.....	19
Gambar 3.2 Alur Proses Penelitian web scraping dan analisis .....	20
Gambar 3.3 Halaman myanimelist.net/people.php .....	26
Gambar 3.4 <i>Output</i> scraping tag href.....	27
Gambar 3.5 Tampilan halaman people .....	28
Gambar 3.6 Data yang diambil di halaman seiyuu .....	30
Gambar 3.7 <i>Output</i> nama karakter anime .....	32
Gambar 3.8 <i>Output</i> jumlah favorite karakter .....	32
Gambar 3.9 Data yang diambil di halaman anime.....	33
Gambar 3.10 <i>Output</i> Scraping halaman anime .....	35
Gambar 3.11 Hasil scraping myanimelist.....	36
Gambar 3.12 <i>Source code</i> dan Tampilan data frame di google collaboratory.....	37
Gambar 3.13 Mengurutkan nama sesuai banyak anime yang diperankan.....	38
Gambar 3.14 Menghapus 67 nama dari data frame .....	39
Gambar 3.15 Tampilan data frame setelah 67 nama dihapus .....	39
Gambar 3.16 Data berdasarkan nama seiyuu dan nama karakter .....	40
Gambar 3.17 hasil akhir pengelompokan data.....	41
Gambar 3.18 Tampilan akhir data frame setelah diproses .....	42
Gambar 3.19 Rumus uji korelasi.....	43
Gambar 3.20 Pembuatan new project di Pycharm.....	44
Gambar 3.21 Struktur folder project .....	45

Gambar 3.22 Tampilan setting .....	45
Gambar 3.23 Tampilan Python Interpreter .....	46
Gambar 3.24 Daftar library yang digunakan.....	46
Gambar 3.25 Pembuatan file data_visualization.py.....	46
Gambar 4.1 Hasil uji korelasi pada website aplikasi streamlit .....	57
Gambar 4.2 Tampilan data setelah penyesuaian .....	59
Gambar 4.3 Tampilan halaman <i>Home</i> .....	60
Gambar 4.4 Tampilan halaman <i>How to use</i> .....	60
Gambar 4.5 Tampilan halaman web aplikasi data visualization.....	61
Gambar 4.6 <i>Bar Chart</i> data jumlah favorite karakter .....	61
Gambar 4.7 <i>Histogram</i> data jumlah favorite karakter .....	62
Gambar 4.8 <i>Scatter plot</i> data jumlah favorite <i>seiyuu</i> dan favorite karakter.....	62
Gambar 4.9 <i>Lineplot</i> data jumlah favorite <i>seiyuu</i> dan favorite karakter.....	63



## DAFTAR KODE

Kode 3.1 <i>Source code</i> universal beautifulsoup .....	18
Kode 3.2 Rancangan data frame pandas .....	19
Kode 3.3 <i>Source code</i> scraping tag href .....	20
Kode 3.4 Merubah <i>source</i> scraping .....	21
Kode 3.5 <i>Source code</i> pembeda seiyuu .....	23
Kode 3.6 <i>Source code</i> scraping halaman seiyuu .....	24
Kode 3.7 <i>Source code</i> scraping halaman anime .....	27
Kode 3.8 <i>Source code</i> pembuatan data frame dan ekspor ke csv .....	29
Kode 3.9 <i>Source code</i> halaman home .....	47
Kode 3.10 <i>Source code</i> halaman <i>how to use</i> .....	47
Kode 3.11 <i>source code</i> halaman data visualization .....	48

