

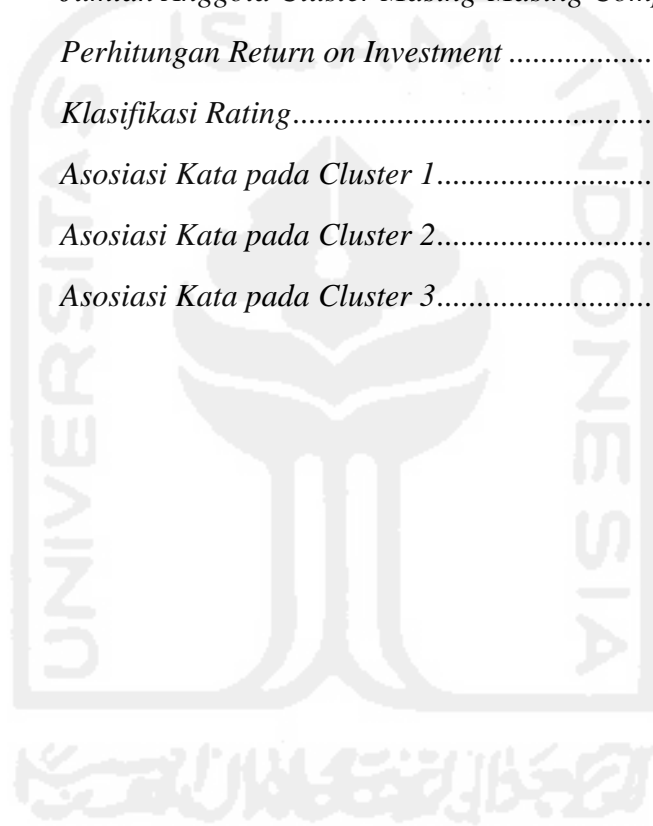
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN PEMBIMBING	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR	iv
DAFTAR ISI	vi
DAFTAR TABEL	viii
DAFTAR GAMBAR	ix
DAFTAR LAMPIRAN	xii
PERNYATAAN	xiii
INTISARI	xiv
ABSTRACT	xv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah	5
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA	7
2.1. Internet Movie Data Base (IMDB).....	7
2.2. Variabel yang terdapat pada Situs IMDB.....	7
2.3. Konten IMDB	8
2.4. Sistem Perangkingan Film pada IMDB.....	11
2.5. The Numbers	12
2.6. Fitur Website The Numbers	13
2.7. Penelitian Terdahulu.....	14
BAB III LANDASAN TEORI	17
3.1. Data Mining.....	17
3.2. Silhouette Coefficient.....	21
3.3. K-Means++.....	22

3.4.	Investasi	28
3.5.	Return on Investment	29
3.6.	Asosiasi Kata	30
BAB IV	METODOLOGI PENELITIAN	31
4.1.	Populasi	31
4.2.	Variabel dan Definisi Operasional Variabel.....	31
4.3.	Sumber Data	33
4.4.	Metode Analisis Data	33
4.5.	Langkah Penelitian	34
BAB V	HASIL DAN PEMBAHASAN	35
5.1.	Scraping Data	35
5.2.	Data Blending.....	41
5.3.	Analisis Deskriptif.....	44
5.4.	Menentukan Jumlah Cluster dengan Silhouette Coefficient	64
5.5.	Analisa K-Means++	66
5.6.	Menghitung ROI (Return on Investment)	73
5.7.	Visualisasi dan Asosiasi	77
BAB VI	PENUTUP	90
6.1.	Kesimpulan	90
6.2.	Saran	91
	DAFTAR PUSTAKA	92
	LAMPIRAN.....	96

DAFTAR TABEL

Tabel 4.1	<i>Definisi Operasional Variabel</i>	31
Tabel 5.1	<i>Silhouette Coefficient</i>	64
Tabel 5.2	<i>Summary Data Masing-Masing Cluster</i>	67
Tabel 5.3	<i>Anggota Masing-Masing Cluster</i>	70
Tabel 5.4	<i>Jumlah Anggota Cluster Masing-Masing Companies</i>	70
Tabel 5.5	<i>Perhitungan Return on Investment</i>	74
Tabel 5.6	<i>Klasifikasi Rating</i>	77
Tabel 5.7	<i>Asosiasi Kata pada Cluster 1</i>	79
Tabel 5.8	<i>Asosiasi Kata pada Cluster 2</i>	84
Tabel 5.9	<i>Asosiasi Kata pada Cluster 3</i>	87



DAFTAR GAMBAR

Gambar 2.1	<i>Konten Review Website IMDB</i>	8
Gambar 2.2	<i>Review Film Website IMDB</i>	9
Gambar 2.3	<i>Review Komentar Film Website IMDB</i>	9
Gambar 2.4	<i>Rating film website IMDB</i>	10
Gambar 2.5	<i>Konten download IMDB</i>	11
Gambar 2.6	<i>Halaman User Rating pada situs IMDB</i>	11
Gambar 2.7	<i>Tampilan Beranda Website The Numbers</i>	13
Gambar 2.8	<i>Tampilan Menu Movie Production Countries</i>	13
Gambar 2.9	<i>Tampilan Menu Movie Budget Website The Number</i>	14
Gambar 3.1	<i>Tahap-Tahap Knowledge Discovery from Data (KDD)</i>	18
Gambar 3.2	<i>Langkah algoritme k-means</i>	25
Gambar 4.1	<i>Flowchart Penelitian</i>	34
Gambar 5.1	<i>Tampilan Website IMDB</i>	35
Gambar 5.2	<i>Menu Advance Search</i>	36
Gambar 5.3	<i>Memilih Kategori Film</i>	36
Gambar 5.4	<i>Tampilan Review Film</i>	37
Gambar 5.5	<i>Aplikasi Webharvy</i>	37
Gambar 5.6	<i>Tampilan Copy URL</i>	38
Gambar 5.7	<i>Seleksi Konten Website</i>	38
Gambar 5.8	<i>Tampilan Captured Data Preview</i>	39
Gambar 5.9	<i>Tampilan Jendela Webharvy Miner</i>	40
Gambar 5.10	<i>Tampilan Website The Numbers</i>	41
Gambar 5.11	<i>Proses Web Scraping pada website The Numbers</i>	41
Gambar 5.12	<i>Tampilan Awal Tableau</i>	42
Gambar 5.13	<i>Tampilan Data dari Website IMDB</i>	42
Gambar 5.14	<i>Tampilan Data dari Website The Numbers</i>	43

Gambar 5.15	<i>Menggabungkan Data dari IMDB dan The Numbers.....</i>	43
Gambar 5.16	<i>Tampilan Data Gabungan IMDB dan The Numbers.....</i>	44
Gambar 5.17	<i>Banyak Film Yang Diproduksi Masing-Masing Companies Tahun 1997-2016</i>	45
Gambar 5.18	<i>Rata-Rata Budget Film Masing-Masing Companies.....</i>	46
Gambar 5.19	<i>Banyaknya Rating Film berdasarkan Kelas Rating.....</i>	47
Gambar 5.20	<i>10 Genre Film dengan Budget Tertinggi</i>	48
Gambar 5.21	<i>10 Genre Film dengan Domestic Gross Tertinggi.....</i>	49
Gambar 5.22	<i>10 Genre Film dengan Worldwide Gross Tertinggi</i>	50
Gambar 5.23	<i>10 Genre Film dengan Domestic Profit Tertinggi.....</i>	51
Gambar 5.24	<i>10 Genre Film dengan Worldwide Profit Tertinggi</i>	52
Gambar 5.19	<i>Rata-Rata Rating Film Tahunan Masing-Masing Companies.....</i>	53
Gambar 5.26	<i>Perbandingan Rata-Rata Tahunan Budget, Domestic Gross, Worldwide Gross, Domestic Profit, Worldwide Profit.</i>	55
Gambar 5.27	<i>Perbandingan Rata-Rata Tahunan Budget, Domestic Gross, Worldwide Gross, Domestic Profit, Worldwide Profit 20 Century Fox.....</i>	56
Gambar 5.28	<i>Perbandingan Rata-Rata Tahunan Budget, Domestic Gross, Worldwide Gross, Domestic Profit, Worldwide Profit Dreamworks.....</i>	57
Gambar 5.29	<i>Perbandingan Rata-Rata Tahunan Budget, Domestic Gross, Worldwide Gross, Domestic Profit, Worldwide Profit MGM.....</i>	58
Gambar 5.30	<i>Perbandingan Rata-Rata Tahunan Budget, Domestic Gross, Worldwide Gross, Domestic Profit, Worldwide Profit Paramount</i>	59
Gambar 5.31	<i>Perbandingan Rata-Rata Tahunan Budget, Domestic Gross, Worldwide Gross, Domestic Profit, Worldwide Profit Sony</i>	60
Gambar 5.32	<i>Perbandingan Rata-Rata Tahunan Budget, Domestic Gross, Worldwide Gross, Domestic Profit, Worldwide Profit Universal</i>	61

Gambar 5.33	<i>Perbandingan rata-rata tahunan budget, domestic gross, worldwide gross, domestic profit, worldwide profit Walt Disney.....</i>	62
Gambar 5.34	<i>Perbandingan Rata-Rata Tahunan Budget, Domestic Gross, Worldwide Gross, Domestic Profit, Worldwide Profit Warner Bros.....</i>	63
Gambar 5.35	<i>Grafik Silhouette Coefficient.....</i>	65
Gambar 5.36	<i>Syntax K-means ++</i>	66
Gambar 5.37	<i>Matrix Cluster K-mean++.....</i>	67
Gambar 5.38	<i>Summary Data</i>	67
Gambar 5.39	<i>Circle Plot.....</i>	68
Gambar 5.40	<i>3 Dimensional Plot</i>	69
Gambar 5.41	<i>Rata-Rata Budget, Domestic Gross dan Worldwide Gross berdasarkan Cluster.....</i>	71
Gambar 5.42	<i>Rata-Rata Domestic Profit dan Worldwide Profit Antar Cluster.....</i>	72
Gambar 5.43	<i>Rata-Rata Rating Antar Cluster.....</i>	73
Gambar 5.44	<i>Rata-Rata ROI Masing-masing Companies.....</i>	75
Gambar 5.45	<i>Rata-Rata ROI berdasarkan Genre</i>	76
Gambar 5.46	<i>Kata yang Sering Muncul pada Cluster 1.....</i>	78
Gambar 5.47	<i>Wordcloud Kata yang Sering Muncul pada Cluster 1.....</i>	78
Gambar 5.48	<i>Kata yang Sering Muncul pada Cluster 2.....</i>	82
Gambar 5.49	<i>Wordcloud Kata yang Sering Muncul pada Cluster 2.....</i>	83
Gambar 5.50	<i>Kata yang Sering Muncul pada Cluster 3.....</i>	86
Gambar 5.51	<i>Wordcloud Kata yang Sering Muncul pada Cluster 3.....</i>	87

DAFTAR LAMPIRAN

- Lampiran 1. Rangkuman Grafik Heat Map
- Lampiran 2. Script Silhouette Coefficient, K-Means++ dan Association
- Lampiran 3. TOP 25 Artis dan Director dengan Rata-Rata ROI Tertinggi

