

DAFTAR ISI

LEMBAR PENGESAHAN	i
LEMBAR PERNYATAAN ORISINALITAS	ii
ABSTRAK	iii
ABSTRACT	iv
KATA PENGANTAR	v
DAFTAR ISI	vi
DAFTAR GAMBAR DAN ILUSTRASI	viii
DAFTAR TABEL	x
DAFTAR LAMPIRAN	xi
Bab I PENDAHULUAN	1
I.1 Latar Belakang	1
I.2 Rumusan Masalah	4
I.3 Tujuan Penelitian.....	4
I.4 Batasan Penelitian	4
I.5 Manfaat Penelitian.....	5
I.6 Sistematika Pelaporan	5
Bab II LANDASAN TEORI	6
II.1 Data <i>Preprocess</i>	6
II.2 Data <i>Profiling</i>	6
II.3 Single-column Profiling	7
II.3.1 Single-column Profiling : <i>Cardinalities</i>	7
II.3.2 Single-column Profiling : <i>Value distribution</i>	8
II.3.3 Single-column Profiling : <i>Patterns</i>	8
II.4 Variasi Arsitektur Teknologi Data Profiling <i>tools</i> berbasis <i>open source</i> ..	9
II.5 Mekanisme Data Profiling	11
II.6 Pemilihan Metode Penelitian	12
Bab III METODOLOGI PENELITIAN	13
III.1 Konseptual Model	13
III.2 Sistematikan Penelitian	14
III.2.1 Fase Inisiasi.....	14

III.2.2	Fase Data Profiling	14
III.2.3	Fase Pengujian	15
Bab IV	ANALISIS DAN DESAIN	16
IV.1	Analisis Pemilihan Algoritma	16
IV.1.1	Algoritma Untuk <i>Cardinalities</i>	16
IV.1.2	Algoritma Untuk <i>Value Distribution</i>	18
IV.1.3	Algoritma Untuk <i>Data Pattern</i>	19
IV.2	Analisis Pemilihan Tools Untuk Data Profiling	20
Bab V	IMPLEMENTASI DAN PENGUJIAN	24
V.1	Implementasi Ruang Lingkup Studi Kasus	24
V.2	Implementasi	25
V.2.1	Pembuatan koneksi Database pada Pentaho Data Integration	25
V.2.2	Konfigurasi <i>Environment Variables</i> pada Pentaho	26
V.2.3	Implementasi logika <i>Cardinalities</i>	27
V.2.4	Implementasi logika <i>Value Distribution</i>	39
V.2.5	Implementasi logika <i>Data Pattern</i>	41
V.3	Pengujian	47
Bab VI	KESIMPULAN DAN SARAN	51
VI.1	Kesimpulan	51
VI.2	Saran	51