

DATA *WAREHOUSE* DAN *DASHBOARD REPORTING* MENGUNAKAN HADOOP DAN R-SHINY DALAM PENERAPAN ARSITEKTUR MEDALLION

MUHAMMAD FAZRI NAHAR



**TEKNOLOGI REKAYASA PERANGKAT LUNAK
SEKOLAH VOKASI
INSTITUT PERTANIAN BOGOR
BOGOR
2024**

PERNYATAAN MENGENAI LAPORAN PROYEK AKHIR DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa laporan proyek akhir dengan judul “Data Warehouse dan Dashboard Reporting Menggunakan Hadoop dan R-Shiny dalam Penerapan Arsitektur Medallion” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir laporan proyek akhir ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2024

Muhammad Fazri Nahar
J0303202170



ABSTRAK

MUHAMMAD FAZRI NAHAR. Data *Warehouse* dan *Dashboard Reporting* menggunakan Hadoop dan R-Shiny dalam Penerapan Arsitektur Medallion. Dibimbing oleh MEDHANITA DEWI RENANTI.

PT XYZ, anak perusahaan XYZ.inc, merupakan produsen mainan terbesar dunia yang fokus pada inovasi berkelanjutan. Kompleksitas pengelolaan sumber daya dan data dalam proses produksi menjadi tantangan signifikan dalam upaya mencapai visi globalnya. Penelitian ini bertujuan mengatasi tantangan tersebut dengan memodifikasi arsitektur medallion dan mengimplementasikan ekosistem Hadoop untuk manajemen data yang lebih efisien. Menggunakan metode *Project Management Knowledge Areas* (PMKA), penelitian ini menghasilkan modifikasi arsitektur medallion yang memungkinkan distribusi data efisien antar perangkat dengan sistem operasi berbeda. Data *warehouse* berbasis Hadoop diimplementasikan pada Windows dan MacOS, memanfaatkan komponen ekosistem Hadoop untuk penyimpanan dan Hive untuk pengelolaan data. Integrasi dengan R-Shiny memfasilitasi visualisasi dan pelaporan data yang efektif. Hasil penelitian ini mendukung pengembangan infrastruktur teknologi informasi di industri mainan global, memungkinkan analisis data yang lebih mendalam dan pengambilan keputusan yang lebih informatif.

Kata kunci: arsitektur medallion, data *warehouse*, ekosistem, hadoop, r-shiny

ABSTRACT

MUHAMMAD FAZRI NAHAR. Data Warehouse and Dashboard Reporting Using Hadoop and R-Shiny in Implementing the Medallion Architecture. Supervised by MEDHANITA DEWI RENANTI.

PT XYZ, a subsidiary of XYZ.inc, stands as the world's largest toy manufacturer with a focus on sustainable innovation. The complexity of resource and data management in the production process poses a significant challenge in achieving its global vision. This research aims to address these challenges by modifying the medallion architecture and implementing the Hadoop ecosystem for more efficient data management. Utilizing the Project Management Knowledge Areas (PMKA) method, this study resulted in a modified medallion architecture that enables efficient data distribution across devices with different operating systems. A Hadoop based data warehouse was implemented on Windows and MacOS, leveraging Hadoop ecosystem components for storage and Hive for data management. Integration with R-Shiny facilitates effective data visualization and reporting. The results of this research support the development of information technology infrastructure in the global toy industry, enabling deeper data analysis and more informed decision making.

Keywords: data warehouse, hadoop ecosystem, medallion architecture, r-shiny



Judul Proyek Akhir : Data *Warehouse* dan *Dashboard Reporting* menggunakan Hadoop dan R-Shiny dalam Penerapan Arsitektur Medallion

Nama : Muhammad Fazri Nahar
NIM : J0303202170

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing :
Medhanita Dewi Renanti, S.Kom, M.Kom.



Diketahui oleh

Ketua Program Studi:
Medhanita Dewi Renanti, S.Kom., M.Kom.
NPI 20180 719830512 2 001

Dekan Sekolah Vokasi:
Dr. Ir. Aceng Hidayat, M.T.
NIP 19660 717199203 1 003



Tanggal Ujian:
5 Agustus 2024

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir ini yang berjudul "Pemanfaatan Ekosistem Hadoop sebagai Data *Warehouse* dan *Dashboard Reporting* Menggunakan R-Shiny dalam Penerapan Arsitektur Medallion". Laporan tugas akhir ini disusun untuk memenuhi persyaratan memperoleh gelar Sarjana Terapan pada program studi Teknologi Rekayasa Perangkat Lunak, Sekolah Vokasi IPB University.

Pada kesempatan ini, penulis menyampaikan terima kasih kepada semua pihak yang telah membantu dalam penyusunan laporan Tugas Akhir ini. Ucapan terima kasih penulis ucapkan kepada Ibu Medhanita Dewi Renanti, S.Kom, M.Kom. selaku dosen pembimbing yang telah memberikan arahan dan bimbingan yang sangat berharga selama proses pembuatan laporan tugas akhir ini. Ucapan terima kasih dan penghargaan juga penulis sampaikan kepada Ibu Dr. Eng. Annisa, S.Kom., M.Kom. selaku dosen moderator dan penguji pada laporan proyek akhir ini.

Selanjutnya, ungkapan terima kasih disampaikan kepada keluarga atas dukungan, doa, dan kasih sayang mereka sepanjang perjalanan ini serta dukungan mereka dalam penulisan laporan tugas akhir ini. Ucapan terima kasih kepada teman-teman yang selalu memberi dukungan dalam pembuatan laporan tugas akhir ini.

Penulis menyadari bahwa laporan tugas akhir ini masih jauh dari kata sempurna, oleh karena itu kritik dan saran yang membangun sangat diharapkan penulis guna menyempurnakan penelitian ini dengan lebih baik. Semoga penelitian ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2024

Muhammad Fazri Nahar



DAFTAR ISI

DAFTAR TABEL	viii
DAFTAR GAMBAR	viii
DAFTAR LAMPIRAN	viii
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat	2
1.5 Ruang Lingkup	3
II METODE	4
2.1 Lokasi dan Waktu Proyek Akhir	4
2.2 Teknik Pengumpulan Data	4
2.3 Prosedur Kerja	4
2.4 Pengujian	10
III HASIL DAN PEMBAHASAN	11
3.1 <i>Exploratory Data Analysis</i>	11
3.2 Transformasi dan Denormalisasi	12
3.3 <i>Data Warehouse</i> Hadoop	13
3.4 <i>Data Pipeline</i>	17
3.5 <i>Dashboard Reporting</i> R-Shiny	21
3.6 Pengujian	29
IV SIMPULAN DAN SARAN	30
4.1 Simpulan	30
4.2 Saran	30
DAFTAR PUSTAKA	31
LAMPIRAN	33
RIWAYAT HIDUP	41

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



DAFTAR TABEL

1	<i>Research questions</i>	7
2	Kebutuhan perangkat lunak	8
3	Deskripsi data	11
4	Daftar <i>query</i> pembuatan eksternal <i>table</i>	16
5	Daftar <i>Jobs</i> transformasi pada Dagster	18
6	Daftar <i>package</i>	23
7	Hasil uji <i>blackbox</i>	29

DAFTAR GAMBAR

1	PMKA <i>phases</i>	5
2	Manajemen data saat ini	6
3	Data <i>product architecture</i>	9
4	Rancangan penerapan arsitektur Medallion	9
5	<i>Entity relationship diagram</i> imajiner	12
6	<i>Entity relationship diagram</i> imajiner Hasil transformasi	13
7	Lokasi folder <i>bronze</i> dan <i>silver</i>	14
8	Folder <i>silver</i>	15
9	Pembuatan eksternal <i>table</i> dt tahap <i>bronze</i>	15
10	Pembuatan eksternal <i>table</i> dt tahap <i>silver</i>	16
11	Pembuatan <i>table view</i> untuk <i>schedule_wo_view</i>	17
12	<i>User interface</i> Dagster	18
13	<i>Jobs</i> untuk memasukkan data <i>silver</i>	20
14	Pembuatan jadwal <i>update</i> data	21
15	Struktur R-Shiny	22
16	Proses penarikan data dari Hive menggunakan R	22
17	<i>Tab OEE summary</i>	24
18	<i>Tab OEE trend</i>	24
19	<i>Tab availability</i>	25
20	<i>Tab performance</i>	25
21	<i>Tab quality</i>	26
22	Contoh filter <i>weekending</i>	27
23	Contoh filter <i>date</i>	27
24	<i>Flowchart month dan weekending</i>	28

DAFTAR LAMPIRAN

1	<i>Query</i> eksternal <i>table</i> <i>downtime_data</i>	35
2	<i>Query</i> eksternal <i>table</i> <i>reject_data</i>	35
3	<i>Query</i> eksternal <i>table</i> <i>dtlog</i>	35
4	<i>Query</i> eksternal <i>table</i> <i>dtlog silver</i>	36
5	<i>Query</i> eksternal <i>table</i> <i>oee</i>	36

6	<i>Query eksternal table oee silver</i>	36
7	<i>Query eksternal table performance</i>	37
8	<i>Query eksternal table performance silver</i>	37
9	<i>Query eksternal table qc</i>	37
10	<i>Query eksternal table qc silver</i>	38
11	<i>Query eksternal table schedule_wo</i>	38
12	<i>Query eksternal table schedule_wo_view silver</i>	38
13	<i>Jobs transformasi data performance</i>	39
14	<i>Jobs transformasi data oee</i>	39
15	<i>Jobs transformasi data dtlog</i>	39
16	<i>Jobs transformasi data dt</i>	39
17	<i>Jobs transformasi data qc</i>	39

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University;