

**ANALISIS DATA KUNJUNGAN WISATA PULAU PARI
DENGAN PYSPARK SQL DAN TABLEAU UNTUK INDUSTRI
PARIWISATA**

PRAKTEK KERJA LAPANGAN



Oleh :

MUHAMMAD ANANDA GIOVANNY RAMADHAN

22082010110

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
Universitas Pembangunan Nasional “Veteran” Jawa Timur**

S U R A B A Y A

2024

LEMBAR PENGESAHAN

Judul : ANALISIS DATA KUNJUNGAN WISATA PULAU PARI DENGAN
PYSPARK SQL DAN TABLEAU UNTUK INDUSTRI PARIWISATA
Oleh : Muhammad Ananda Giovanny Ramadhan NPM: 22082010110

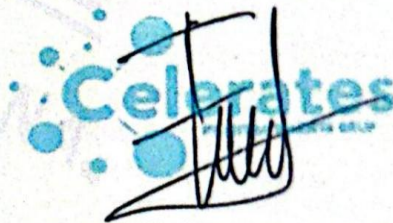
Menyetujui,

Dosen Pembimbing



Prasasti Karunia Farista Ananto, M.IM
NIP.19970704 202406 2 001

Mentor PT Mitra Talenta Grup

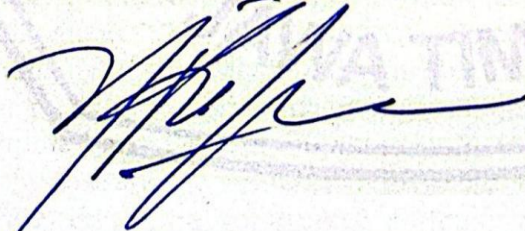


Muhammad Falsal Irsyad
NIP. MTG-CS/062024115

Mengetahui

Dekan

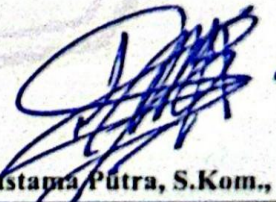
Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, M.T
NIP. 19681126 199403 2 001

Koordinator

Program Studi Sistem Informasi



Agung Brastama Putra, S.Kom., M.Kom
NIP. 19851124 202121 1 003

ABSTRAK

Praktik Kerja Lapangan (PKL) ini dilakukan dengan tujuan untuk mengelola dan menganalisis data kunjungan wisata di Pulau Pari, salah satu destinasi unggulan di Kepulauan Seribu, DKI Jakarta. Dalam era digital, pengelolaan data telah menjadi aspek krusial untuk mendukung pengambilan keputusan berbasis data, terutama dalam sektor pariwisata yang sangat bergantung pada tren kunjungan, preferensi wisatawan, dan strategi pemasaran yang efektif. Penelitian ini menggunakan pendekatan data engineering dengan memanfaatkan teknologi modern seperti Pentaho untuk ingest data, Apache Airflow untuk orkestrasi pipeline, Apache Spark untuk pemrosesan data skala besar, Hadoop untuk penyimpanan data terdistribusi, dan Tableau untuk visualisasi data.

Pipeline data yang dirancang meliputi pengumpulan data dari berbagai sumber, transformasi data menjadi format yang terstruktur, penyimpanan data dalam data warehouse berbasis PostgreSQL dan Hadoop, hingga visualisasi data dalam bentuk dashboard interaktif. Proses ini memungkinkan analisis mendalam terhadap pola kunjungan wisatawan, metode pembayaran yang digunakan, aktivitas yang paling diminati, serta preferensi wisatawan berdasarkan ulasan mereka. Hasil visualisasi menunjukkan bahwa total kunjungan wisatawan mencapai 13.182 orang, dengan aktivitas Tur Mangrove memiliki rating tertinggi dan Bersepeda memiliki rating terendah. Preferensi wisatawan menunjukkan bahwa BBQ menjadi aktivitas paling diminati, sementara metode pembayaran tunai mendominasi transaksi dengan persentase 35%.

KATA PENGANTAR

Puji syukur dipanjatkan kepada Tuhan Yang Maha Esa atas rahmat, hidayah, dan karunia-Nya, yang memungkinkan laporan ini dapat diselesaikan dengan baik. Ucapan terima kasih yang mendalam juga ditujukan kepada Universitas Pembangunan Nasional "Veteran" Jawa Timur atas dukungan dan kesempatan yang diberikan selama pelaksanaan Praktek Kerja Lapangan (PKL) di PT. Mitra Talenta Group.

Laporan ini menyajikan hasil dari proyek pengelolaan data kunjungan wisata Pulau Pari menggunakan pendekatan data engineering dengan teknologi modern seperti Pentaho, Apache Airflow, Apache Spark, Hadoop, dan Tableau. Tujuan utama dari proyek ini adalah merancang dan mengimplementasikan pipeline data yang terintegrasi untuk mendukung pengelolaan data besar serta memberikan wawasan berbasis data kepada pengelola wisata Pulau Pari. Laporan ini menguraikan berbagai aspek proyek, termasuk latar belakang, tujuan, metodologi, hasil analisis, serta manfaat yang diharapkan. Proyek ini dirancang untuk membantu pengelola wisata dalam mengambil keputusan strategis yang lebih efektif, efisien, dan berkelanjutan.

Penyusunan laporan ini tidak terlepas dari dukungan dan bimbingan berbagai pihak. Ucapan terima kasih disampaikan kepada:

1. Tuhan Yang Maha Esa, atas petunjuk dan rahmat-Nya yang telah memudahkan proses penyusunan laporan ini.
2. Kedua orang tua dan keluarga yang terus memberikan dukungan dan motivasi.
3. **Ibu Prasasti Karunia Farista Ananto, S.Kom., M.Kom.**, selaku Dosen Pembimbing Praktek Kerja Lapangan Prodi Sistem Informasi Universitas Pembangunan Nasional "Veteran" Jawa Timur, yang telah memberikan izin, bimbingan, dan dukungan selama proses Praktek Kerja Lapangan berlangsung.

4. **Bapak Muhlis Tahir, S.pd., M.Tr.Kom** selaku Pembimbing Lapangan (PKL) di PT Mitra Talenta Grup (Celerates), yang telah memberikan arahan, bimbingan, dan dukungan selama proses PKL berlangsung.
5. Para mentor di PT Mitra Talenta Grup (Celerates) yang telah memberikan bimbingan dan dukungan teknis selama pelaksanaan proyek ini.
6. Teman-teman Prodi Sistem Informasi yang selalu memberikan dukungan dan motivasi.

Laporan ini telah disusun dengan upaya dan dedikasi penuh. Namun, penulis menyadari masih terdapat kekurangan dalam laporan ini, sehingga saran dan kritik yang membangun sangat diharapkan untuk perbaikan di masa mendatang. Semoga laporan ini dapat memberikan manfaat dan wawasan baru bagi para pembaca.

Surabaya, 6 Januari 2025

Penyusun

DAFTAR ISI

ABSTRAK	ii
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR TABEL	vii
DAFTAR GAMBAR	viii
DAFTAR LAMPIRAN	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan PKL	4
1.4 Manfaat	5
BAB II GAMBARAN UMUM TEMPAT PKL	8
2.1 Profil Organisasi	8
2.2 Struktur Organisasi	10
2.3 Visi Dan Misi Organisasi	10
2.3.1 Visi Perusahaan	10
2.3.2 Misi Perusahaan	11
BAB III PELAKSANAAN PKL	12
3.1 Tinjauan Pustaka	12
3.1.1 Data Engineering	12
3.1.2 Data Ingest	13
3.1.3 Data Warehouse	14
3.1.4 Data Governance	16
3.1.5 Visualisasi Data	18
3.1.6 Tableau	19
3.1.7 Hadoop	19
3.1.8 Pentaho Data Integration (PDI)	19
3.1.9 Pengertian Apache Spark	20
3.1.10 Apache Airflow	20
3.1.11 PostgreSQL	20

3.2	Waktu Dan Pelaksanaan PKL	21
3.2.1	Waktu Dan Tempat Penelitian	21
3.2.2	Pelaksanaan	21
BAB IV HASIL DAN PEMBAHASAN		25
4.1	Metode Pelaksanaan Praktek Kerja Lapangan	25
4.2	Analisis Kebutuhan	26
4.3	Perancangan	28
4.3.1	Flow Diagram	28
4.4	Implementasi	30
4.4.1	Membuat Raw Data	30
4.4.2	Transformasi Pentaho	35
4.4.3	Pyspark	34
4.4.4	Airflow	46
4.4.5	Data Visualitazion	48
BAB V PENUTUP		58
5.1	Kesimpulan	58
5.2	Saran	59
DAFTAR PUSTAKA		60
LAMPIRAN		61
Lampiran 1.	Surat Keterangan Rahasia	62
Lampiran 2.	Lembar Pengesahan	63

DAFTAR TABEL

Tabel 1 Jadwal Kegiatan	21
--------------------------------	-----------

DAFTAR GAMBAR

Gambar 2. 1 Logo PT Mitra Talenta Grup (Celerates)	8
Gambar 4. 1 Diagram Alir PKL	25
Gambar 4. 2 Flow Diagram	28
Gambar 4. 3 Library dan Inisialisasi	31
Gambar 4. 4 Fungsi Untuk Data	31
Gambar 4. 5 Data Master	32
Gambar 4. 6 Generate Data	33
Gambar 4. 7 Konversi Ke DataFrame	33
Gambar 4. 8 Penyesuaian Jumlah Orang Berdasarkan ID Wisatawan	34
Gambar 4. 9 Penyimpanan Ke CSV	34
Gambar 4. 10 Output	34
Gambar 4. 11 Ingest Data	35
Gambar 4. 12 Select Values	35
Gambar 4. 13 Insert Update	36
Gambar 4. 14 Inisialisasi Library	36
Gambar 4. 15 Membaca Tabel Dari PostgreSQL	37
Gambar 4. 16 Validasi Schema	37
Gambar 4. 17 Validasi Hasil Join	38
Gambar 4. 18 Output	38
Gambar 4. 19 Inisialisasi dan koneksi ke PostgreSQL	39
Gambar 4. 20 Membaca Tabel Utama Dari PostgreSQL	40
Gambar 4. 21 Validasi Schema	41
Gambar 4. 22 Hasil Join	41
Gambar 4. 23 Output	42
Gambar 4. 24 Inisialisasi	43
Gambar 4. 25 Memuat Tabel	43
Gambar 4. 26 Melakukan Join	44
Gambar 4. 27 Validasi Hasil Join	44

Gambar 4. 28 Output Hasil Join	45
Gambar 4. 29 Inisialisasi koneksi	45
Gambar 4. 30 Memuat Data Ke PostgreSQL	46
Gambar 4. 31 Menghitung data invalid	46
Gambar 4. 32 Inisialisasi Ke PostgreSQL	47
Gambar 4. 33 Inisialisasi Ke PostgreSQL	47
Gambar 4. 34 Hasil Validasi	47
Gambar 4. 35 Inisialisasi	48
Gambar 4. 36 Koneksi Ke PostgreSQL	49
Gambar 4. 37 Memuat Data Dari PostgreSQL	49
Gambar 4. 38 Menentukan Path HDFS	49
Gambar 4. 39 Membaca Kembali File CSV	50
Gambar 4. 40 Dashboard Tableau	50
Gambar 4. 41 Instalasi Library	51
Gambar 4. 42 Parameter	52
Gambar 4. 43 Definisi DAG	52
Gambar 4. 44 Masking Data	53
Gambar 4. 45 Simpan Ke PostgreSQL	53
Gambar 4. 46 Validasi Dan Backup	54
Gambar 4. 47 Output	55
Gambar 4. 48 Airflow	57

DAFTAR LAMPIRAN

Lampiran 1 Surat Pernyataan Rahasia	61
--	-----------