

LAPORAN AKHIR
STUDI INDEPENDEN BERSERTIFIKAT
DATA ANALYTICS
DI GREAT EDU GLOBAL MAHARDIKA

Diajukan untuk memenuhi persyaratan kelulusan
Program MSIB MBKM

oleh :

Muhammad Anwar Lirboyo / 201910225319



PROGRAM STUDI INFORMATIKA
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2022

Lembar Pengesahan

Program Studi Informatika Universitas Bhayangkara Jakarta Raya

STUDI INDEPENDEN BERSERTIFIKAT DATA ANALYTICS

Di Great Edu (PT. Great Edu Mahardika)

oleh :

Muhammad Anwar Lirboyo / 201910225319

disetujui dan disahkan sebagai

Laporan Magang atau Studi Independen Bersertifikat Kampus Merdeka

Bekasi, 21 Desember 2022

Pembimbing Magang atau Studi Independen Informatika & Universitas
Bhayangkara Jakarta Raya



Adi Muhajirin, S.Kom., M.Kom., M.M.

NIP: 021503028

Lembar Pengesahan

STUDI INDEPENDEN BERSERTIFIKAT DATA ANALYTICS

Di Great Edu (PT. Great Edu Global Mahardika)

oleh :

Muhammad Anwar Lirboyo / 201910225319

disetujui dan disahkan sebagai

Laporan Magang atau Studi Independen Bersertifikat Kampus Merdeka

Jakarta, 28 Desember 2022

Partnership Officer



Robert Edy Sudarwan

Abstraksi

Studi Independen Bersertifikat adalah bagian dari program Kampus Merdeka yang bertujuan untuk memberikan kesempatan kepada mahasiswa untuk belajar dan mengembangkan diri melalui aktivitas di luar kelas perkuliahan, namun tetap diakui sebagai bagian dari perkuliahan. Laporan ini memuat laporan Studi Independen Bersertifikat yang diselenggarakan oleh PT GreatEdu Global Mahardika dengan topik pembelajaran Data Analytics. GreatEdu adalah perusahaan teknologi Pendidikan asal Indonesia yang berbasis digital platform. Di mana proses pembelajarannya memanfaatkan beberapa digital platform, seperti Zoom Meeting, Google Meet, dan WhatsApp Group sebagai tempat berkoordinasi antara pihak GreatEdu dan mahasiswa. Pembelajaran berupa live class di setiap minggunya yang terbagi dalam beberapa sesi, pengerjaan tugas, dan sesi mentoring. Pada akhir periode program, peserta dibentuk menjadi beberapa kelompok untuk pengerjaan final project. Dalam proses pengerjaannya, terdapat enam minggu sesi mentoring yang memuat proses dari persiapan pengerjaan hingga tahap finalisasi dan pengumpulan final project. Final project yang dibahas pada laporan ini berjudul “Analisis Sentimen Lazada Indonesia Review”. Secara keseluruhan, proses pembelajaran di GreatEdu dapat membekali mahasiswa yang ingin mengenal dunia data analytics. Selain itu, dengan adanya final project juga bermanfaat bagi mahasiswa untuk praktik langsung mengerjakan suatu proyek data analisis, sehingga pada akhirnya mahasiswa memiliki portofolio.

Kata kunci: Studi Independen Bersertifikat, GreatEdu, Data Analytics, Final Project

Kata Pengantar

Puji syukur kami panjatkan kepada Tuhan Yang Maha Esa karena atas limpahan Rahmat-Nya kami dapat menyelesaikan laporan akhir Studi Independen Data Analytics DI Great Edu Global Mahardika. Selain itu, atas segala bentuk dukungan yang diberikan dalam penyusunan laporan ini, penulis juga mengucapkan terima kasih kepada:

1. Orang tua dan keluarga yang senantiasa mendoakan dan memberikan dukungan.
2. Bapak Adi Muhajirin, S.kom., M.Kom., M.M. selaku Dosen Pembimbing Lapangan Studi Independen Bersertifikat.
3. *Coach* Jashinta Kresna Senja, selaku *Homeroom* kelas Data Analytics.
4. PT Great Edu Global Mahardika sebagai perusahaan yang telah menjadi tempat Studi Independen Bersertifikat.
5. Teman-teman kelompok Team 4tune, serta.
6. Seluruh pihak yang tidak dapat disebutkan satu persatu yang telah mendukung dalam penyelesaian laporan ini.

Penulis menyadari bahwa laporan ini masih jauh dari kesempurnaan karena pengetahuan dan pengalaman kami yang masih sangat terbatas. Penulis sangat mengharapkan kritik dan saran yang bersifat membangun dari berbagai pihak dan penulis berharap dengan adanya laporan ini, kami dapat menambah wawasan, pengetahuan serta pengalaman bagi penulis.

Bekasi, 28 Desember 2022



Muhammad Anwar Lirboyo
NPM. 201910225319

Daftar Isi

Bab I	Pendahuluan	1
I.1	Latar belakang	1
I.2	Lingkup	2
I.3	Tujuan	3
Bab II	Great Edu Global Mahardika	1
II.1	Struktur Organisasi	1
II.2	Lingkup Pekerjaan	6
II.3	Deskripsi Pekerjaan	6
II.4	Jadwal Kerja	7
Bab III	Project Akhir “Analisis Sentimen Lazada Indonesia Reviews”	1
III.1	Latar Belakang Proyek Akhir	1
III.2	Proses Pelaksanaan Proyek Akhir	1
III.3	Hasil Proyek Akhir	8
Bab IV	Penutup	1
IV.1	Kesimpulan	1
IV.2	Saran	1
Bab V	Referensi	vii
Bab VI	Lampiran A. TOR	1
Bab VII	Lampiran B. Log Activity	1
Bab VIII	Lampiran C. Dokumen Teknik	1

Daftar Tabel

Tabel 2.1 Agenda Kelas	7
Tabel 2.2 Jadwal Kegiatan Pembelajaran	8
Tabel 2.3 Jadwal Pelaksanaan Project Akhir	10

Daftar Gambar

Gambar 2.1 Logo Great Edu	1
Gambar 2.2 Struktur Organisasi Great Edu	2
Gambar 3.1 Dataset Reviews Lazada Indonesia	1
Gambar 3.2 Dataset Items Lazada Indonesia	2
Gambar 3.3. Mengimport library di Google Colab	3
Gambar 3.4 Input dataset yang digunakan	3
Gambar 3.5. Mengecek missing value	4
Gambar 3.6. Mengecek missing value	4
Gambar 3.7. review paling banyak	5
Gambar 3.8. diagram review paling banyak	5
Gambar 3.9. melihat category dengan like paling banyak	5
Gambar 3.10. diagram category dengan like paling banyak	5
Gambar 3.11. melihat category dengan upvote paling banyak	6
Gambar 3.12. melihat category dengan upvote paling banyak	6
Gambar 3.13. Komposisi items menurut jumlah kategorinya	6
Gambar 3.14. Korelasi jumlah transaksi harian dan nilai transaksi	7
Gambar 3.15. WordCloud untuk visualisasi data teks	8

Bab I Pendahuluan

I.1 Latar belakang

Kampus Merdeka merupakan bagian dari kebijakan Merdeka Belajar yang diselenggarakan oleh Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi Republik Indonesia yang memberikan kesempatan kepada mahasiswa di Indonesia untuk mengasah kemampuan dan juga bakat sebagai bekal persiapan untuk terjun ke dunia profesional. Program kampus merdeka menyediakan beberapa program yang dapat diikuti oleh mahasiswa, salah satunya adalah program Magang dan Studi Independen Bersertifikat (MSIB). MSIB merupakan gabungan dari dua program, yaitu program magang dan program studi independen. Pada program studi independen, mahasiswa difasilitasi dengan berbagai bidang keilmuan yang disertai dengan pilihan mitra yang berkualitas.

Salah satu mitra yang ada di Studi Independen Bersertifikat Kampus Merdeka yaitu PT. GreatEdu Global Mahardika. GreatEdu merupakan platform edukasi yang berkonsentrasi pada peningkatan *hard skill* dan juga *soft skill* khususnya *digital skill* seperti *Digital Marketing*, *UI/UX*, Ilmu data dan lainnya. Di era digital saat ini, kebutuhan sumber daya manusia di bidang data banyak dicari oleh perusahaan, terutama perusahaan digital. Banyak sekali profesi baru yang menjadi komoditas utama dalam hal tersebut, seperti *data scientist*, *data analyst*, *data engineer*, dan *artificial intelligence*.

Profesi *data analyst* merupakan profesi dengan permintaan yang cukup banyak pada saat ini hingga beberapa tahun kedepan. Mengutip dari platform *job seeker* di Indonesia, dibutuhkan lebih dari 500 profesi sebagai *data analyst* setiap bulannya. Karena belum adanya sekolah khusus data, maka perlu ada pelatihan bagi calon *data analyst* untuk meningkatkan keterampilan yang dibutuhkan. *Data analyst* merupakan profesi yang secara umum bertugas menganalisis data perusahaan untuk menghasilkan informasi dari data yang ada sehingga bisa diterapkan dalam keputusan bisnis. Pentingnya peran *data analyst* dalam sebuah bisnis perusahaan, ini membuat permintaan terhadap seorang *data analyst* semakin tinggi sehingga menarik para awam untuk mempelajari *data analytics*.

Program Studi Independen Bersertifikat *Data Analytics* terdiri dari kegiatan pembelajaran materi per individu dengan melalui cara *synchronous* dan *asynchronous*. Alur materi mengacu pada proses kerja Data Analisis di industri yang menggunakan pendekatan *design thinking*. Mahasiswa akan merancang solusi produk melalui lima proses *iteratif* yaitu *empathize*, *define*, *ideate*, *prototype*, dan *test*.

Pada program studi independen *Data Analytics* yang diselenggarakan oleh GreatEdu, peserta difasilitasi dengan berbagai kelas pembelajaran dari materi dasar hingga materi lanjutan mengenai *data analytics*. Menjelang akhir periode pelaksanaan program selesai, peserta akan mendapatkan penugasan yang berupa proyek akhir. Di mana dalam pengerjaan proyek akhir, mahasiswa akan mengimplementasikan apa yang telah diperoleh dari kelas-kelas pembelajaran mengenai *data analytics*.

I.2 Lingkup

A. Great Edu Data Analytics

Mahasiswa mendapatkan pembelajaran materi per individu dengan melalui cara *synchronous* dan *asynchronous*; Alur materi mengacu pada proses kerja Data Analisis di industri yang menggunakan pendekatan *design thinking*. Peserta akan menganalisis data membuat deck rekomendasi dari hasil analisis data dengan metode Crisp-Dm.

B. CRISP-DM

Adalah The Cross Industry Standard Process for Data Mining merupakan standard yang dikembangkan sejak tahun 1996 di Eropa. CRISP DM terdiri dari 6 tahap yaitu: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluasi, Deployment.

C. Machine Learning

Secara definisi, machine learning adalah cabang dari ilmu kecerdasan buatan yang berfokus pada pembangunan sebuah sistem agar mampu

belajar dari data-data yang diperolehnya. Menurut Arthur Samuel, machine learning adalah bidang studi yang mempelajari kemampuan program komputer untuk belajar tanpa secara eksplisit diprogram. Implementasi dari Machine Learning ini membutuhkan data. Tanpa adanya data Machine Learning tidak dapat bekerja. Data yang ada yang biasanya dibagi menjadi dua, yaitu data training dan data testing. Data training digunakan untuk melatih algoritma, sedangkan data testing digunakan untuk mengetahui performa algoritma yang telah dilatih sebelumnya ketika menemukan data baru yang belum pernah dilihat. Mahasiswa mendapatkan materi sekaligus praktikum langsung dalam pembelajaran Machine Learning mulai dari basic hingga mahasiswa bisa menciptakan sebuah projek .

D. Python

Python adalah bahasa pemrograman tujuan umum yang ditafsirkan, tingkat tinggi. Dibuat oleh Guido Van Rossum dan pertama kali dirilis pada tahun 1991, filosofi desain python menekankan keterbacaan kode. Konstruksi bahasanya dan pendekatan berorientasi objek bertujuan untuk membantu pemrograman menulis kode yang jelas dan logis untuk proyek skala kecil dan besar.

E. Berpikir Kritis (Critical Thinking)

Mahasiswa belajar bagaimana berpikir kritis mampu menyimpulkan dari apa yang diketahuinya, dan mengetahui cara memanfaatkan informasi untuk memecahkan masalah, dan mencari sumber-sumber informasi yang relevan untuk dirinya.

I.3 Tujuan

Studi Independen Bersertifikat merupakan bagian dari program Kampus Merdeka yang bertujuan untuk memberikan kesempatan kepada para mahasiswa

untuk belajar dan mengembangkan diri melalui aktivitas di luar kelas perkuliahan, namun tetap diakui sebagai bagian dari perkuliahan.

Pada program Studi Independen Data Analytics yang diselenggarakan oleh Great Edu Global Mahardika, mahasiswa diharapkan :

1. Memiliki wawasan mengenai profesi Data Analysts
2. Memahami softskills yang dibutuhkan seorang Data Analysts
3. Memahami Tools dan Teknik yang digunakan dalam Data Analytics
4. Dapat mempraktikkan Tools dan Teknik yang sudah dipelajari
5. Membuat deck rekomendasi dari hasil analisis data dengan metode CRISP DM pada final project secara berkelompok.

Bab II Great Edu Global Mahardika

II.1 Struktur Organisasi



Gambar 2.1 Logo Great Edu

GreatEdu merupakan platform belajar online untuk upskilling dan reskilling. Ratusan pengajar profesional telah bergabung di GreatEdu di tahun 2020.

GreatEdu menghadirkan Workshop dan Video belajar yang berkonsentrasi untuk meningkatkan softskill dan hardskill yang dibutuhkan dalam pekerjaan dan pengembangan karir. Beragam kategori workshop kami sediakan seperti UI/UX, Data, Digital Marketing, Career Development, Food Photography, dan lainnya.

Visi:

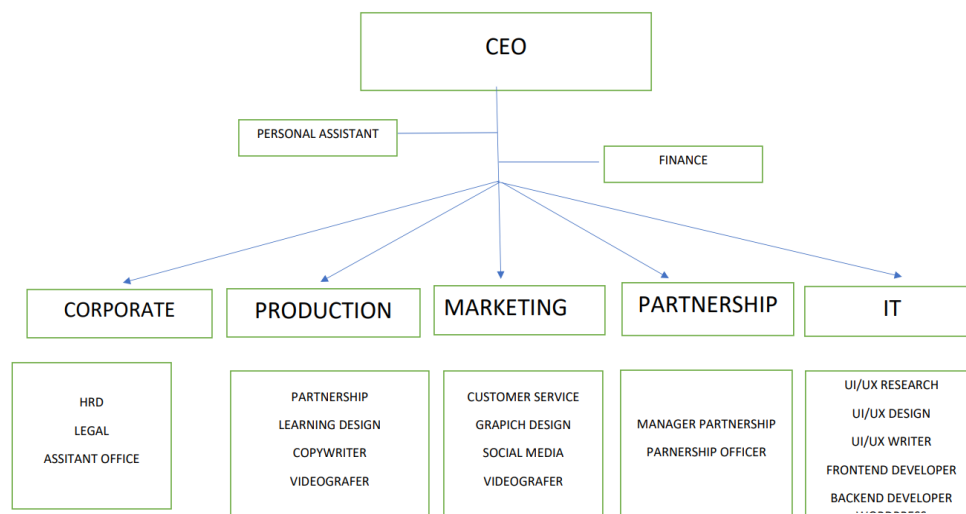
Memberikan pembelajaran berbasis keterampilan transformatif terbaik untuk para pencari kerja & pencipta lapangan kerja.

Misi:

1. Membangun jaringan nasional untuk menyampaikan kurikulum keterampilan masa depan berbasis sertifikasi melalui Platform Konten Digital.
2. Secara proaktif bekerja dengan pemerintah & organisasi dengan mengubah tenaga kerja mereka agar sesuai dengan perubahan pekerjaan yang terjadi karena Industri 4.0.

3. Melatih pemuda dengan keterampilan kewirausahaan & mencocokkan mereka dengan peluang masa depan yang muncul di berbagai industri.
4. Menghubungkan jaringan inkubator dan akselerator yang dikurasi ke industri, investor, dan ekosistem start-up global.

Struktur organisasi Great Edu dapat dilihat pada Gambar 2.2.



Gambar 2.2 Struktur Organisasi Great Edu

- a. **CEO:** Chief Executive Officer (CEO) adalah orang yang bertanggung jawab penuh atas jalannya perusahaan yang dipimpinnya. CEO adalah orang yang bertindak sebagai perantara antara manajer dan direktur atau investor. Beberapa tugas CEO yaitu memimpin perusahaan, merencanakan strategi bisnis, dan mengevaluasi para eksekutif.
- b. **Personal Assistant:** Personal assistant adalah seseorang yang memiliki tugas untuk membantu orang tertentu seperti atasan dalam melakukan aktivitasnya mulai dari hal yang bersifat pribadi hingga pekerjaan.

- c. **Finance:** Tugas yang ditangani oleh tim keuangan meliputi pencatatan transaksi, pengelolaan pendapatan dan arus kas, serta menjalankan fungsi akuntansi, termasuk pembayaran. Departemen keuangan harus bertanggung jawab penuh atas arus informasi keuangan. Tingkat tanggung jawab tergantung pada bidang pekerjaan perusahaan dan kompleksitas perusahaan.
- d. **HRD:** Human Resource Development (HRD) merupakan posisi sentral dan tanggung jawab besar dalam perusahaan. Tanggung jawabnya berkisar dari menerapkan kebijakan perusahaan, peraturan dan strategi untuk menjaga hubungan karyawan yang baik antara karyawan dan perusahaan.
- e. **Legal:** Legal Officer bertanggung jawab untuk memberikan pengarahan kepada direksi mengenai masalah hukum, potensi risiko, dan tindakan apa yang perlu diambil ketika terjadi suatu masalah.
- f. **Assistant Office:** Office Assistant menangani tugas-tugas organisasi dan administratif. Peran mereka yaitu melibatkan mengatur file, menjadwalkan rapat dan mengelola kalender, menulis dan mengoreksi, mengirim email, memelihara persediaan, menyambut tamu, dan banyak lagi.
- g. **Partnership:** Partnership Officer akan bertanggung jawab membangun jaringan kolaboratif dengan berbagai pemangku kepentingan dan menjaga hubungan baik dengan mitra untuk menciptakan keberlanjutan.
- h. **Learning Design:** Learning Designer berperan penting dalam merancang kurikulum. Mereka menyiapkan sumber daya penting untuk membantu membimbing siswa melalui proses pembelajaran yang mudah diakses berdasarkan penelitian dan teori.
- i. **Copywriter:** Tugas seorang copywriter adalah menulis sebuah teks promosi. Tapi tidak hanya itu. beberapa tugas lainnya yaitu

brainstorming dan membuat konsep periklanan dan kampanye pemasaran untuk produk atau layanan yang ditawarkan.

- j. **Videografer:** Tugas Videografer adalah membuat konsep, merekam, dan mengedit video untuk iklan dan media sosial.
- k. **Customer Service:** Tugas Customer Service (CS) pada hakekatnya adalah menjadi jembatan antara pelanggan dengan perusahaan. CS adalah citra perusahaan tempat dia bekerja. Oleh karena itu, CS berkewajiban menjaga nama baik perusahaan. Ini adalah tugas CS. Beberapa tugas CS yaitu membangun hubungan yang baik dengan pelanggan, melakukan penanganan terhadap keluhan – keluhan yang disampaikan konsumen, dan memahami produk/jasa yang ditawarkan.
- l. **Graphic Design:** Graphic Designer bertugas menciptakan ilustrasi, tipografi, fotografi atau gambar bergerak baik untuk media penerbitan, cetak, maupun elektronik. Graphic Designer bertanggung jawab atas tampilan materi promosi produk. Tugasnya menyampaikan informasi tentang produk dengan cara yang menarik, dengan memperhatikan keinginan pelanggan.
- m. **Manager Partnership:** Partnership Manager bertanggung jawab memimpin Departemen Partnership. Posisi ini akan mengkoordinasikan semua divisi dalam Departemen Partnership agar berjalan secara efektif dan efisien dalam melaksanakan tugas.
- n. **Partnership Officer:** Partnership Officer bertanggung jawab untuk menganalisis dan memetakan calon mitra, mengembangkan proposal kemitraan, dan mengelola daftar mitra suatu perusahaan.
- o. **Social Media:** Social Media Specialist bertugas merancang dan membuat rencana media sosial yang selaras dengan merek atau identitas perusahaan. Kemudian juga mengembangkan strategi media sosial melalui penelitian media sosial, perpesanan, dan pemahaman audiens.

- p. **UI/UX Research:** UI/UX Researcher berfokus pada riset pasar sasaran dan mengenali keinginan Anda untuk memberikan wawasan realistis ke dalam keseluruhan proses desain.
- q. **UI/UX Design:** UI/UX Designer bertanggung jawab untuk menentukan tampilan dan nuansa aplikasi dan/atau situs web. Jadi ketika seorang desainer UX mendefinisikan bagaimana sebuah aplikasi dan/atau situs web dapat bekerja dengan mudah. Namun keduanya harus mendasarkan pekerjaannya pada hasil riset, agar aplikasi atau website yang direncanakan benar-benar efektif.
- r. **UI/UX Writer:** UI/UX Writer bertugas membuat salinan untuk aplikasi, situs web, dan produk digital lainnya. Profesi ini bertugas menemukan kata-kata terbaik di menu, tombol, label, dan chatbot untuk membantu pengguna menavigasi produk dengan mudah.
- s. **Frontend Developer:** Frontend Developer adalah pengembang yang terkait erat dengan pengembang UI/UX. Karena tugas umum front end developer adalah mewujudkan hasil akhir dari pengembangan perangkat lunak berdasarkan desain dari desainer UI/UX. Pengembang frontend bertanggung jawab untuk merancang dan membangun aplikasi sisi klien untuk memastikan tampilan aplikasi yang ramah pengguna.
- t. **Backend Developer:** Tugas frontend developer tentunya adalah membuat layout website yang nantinya bisa langsung dilihat oleh pengguna. Berbeda dengan sistem backend yang membuat sistem tidak terlihat oleh pengguna (di balik layar). Berikut ini merupakan beberapa tugas yang dimiliki oleh backend developer yaitu merancang struktur model data, menulis kode aplikasi untuk meningkatkan keamanan, dan memecahkan masalah sisi server.

II.2 Lingkup Pekerjaan

Sebagai mitra Kampus Merdeka MSIB, GreatEdu menyediakan 2 program pelatihan berbeda. Peserta diminta memilih satu dari program yang tersedia, peserta akan menjalani *bootcamp* selama 5 bulan untuk meningkatkan kemampuan teknis dan persiapan mengerjakan final project. GreatEdu mewajibkan seluruh peserta untuk menyelesaikan komponen pembelajaran yang diinstruksikan yaitu, pembelajaran mingguan/*weekly assignments*, quiz, *mentoring*, dan *final project*. Dan semua materi pembelajaran dapat di akses di <https://greatedu.co.id> Adapun sistem pembelajaran di kurikulum ini terdiri atas:

- a. *Lecture Session*
Lecture session adalah sesi belajar-mengajar yang dilakukan dalam kelas secara online bersama expert di bidangnya, pada waktu yang telah disepakati. *Lecture session* wajib diikuti oleh seluruh peserta. Peserta yang datang terlambat harus mengejar ketertinggalan dengan mengetahui kompetensi yang diajarkan, dan melakukan eksplorasi terhadap kompetensi tersebut.
- b. *Mentoring Session*
Mentoring session adalah sesi tanya jawab yang dilakukan antara mentor dan siswa secara online pada waktu yang telah disepakati di luar jam belajar mengajar. *Mentoring session* dilakukan setiap akhir pekan setiap minggunya dengan durasi per sesi maksimum dua (2) jam yang wajib diikuti oleh seluruh siswa.
- c. *Practice Session*
Practice session adalah sesi praktek dimana instruktur akan memberikan sejumlah tugas untuk dikerjakan oleh peserta. Selama *practice session*, peserta wajib mengerjakan seluruh *assignment* yang diberikan dengan penuh tanggung jawab dan dapat meminta bantuan instruktur jika diperlukan.

II.3 Deskripsi Pekerjaan

Berikut adalah deskripsi pekerjaan student sebelum pengerjaan PA:

- a. Mengikuti *lecture session* pada weekdays dengan jadwal yang sudah ditentukan oleh pihak GreatEdu pada awal minggu pertemuan.
- b. Mengikuti *live quiz session* pada saat pembelajaran berlangsung dengan expert
- c. Mengulang materi yang telah disampaikan pada saat *lecture session* (self-study).
- d. Mengerjakan *assignment* yang diberikan di *google classroom* oleh expert hingga batas waktu tertentu.

- e. Mengerjakan mini project (practice session) yang diberikan di google classroom oleh expert hingga batas waktu tertentu.
- f. Mengikuti Mentoring Session (review weekly) pada hari sabtu, pukul 10.00-11.30 WIB.
- g. Peserta memiliki peran Data Engineer dan Data Analyst selama pengerjaan final project, dengan deskripsi pekerjaan sebagai berikut:
- h. Mencari referensi code dan jurnal yang tepat untuk mengolah data yang telah diberikan oleh mentor.
- i. Melakukan analisis data berawal dan pengolahan data seperti menangani data kosong (missing value) dan mengelola deskripsi data untuk dijadikan visualisasi
- j. Melakukan tahap Exploration Data Analysis (EDA) untuk membuat visualisasi data sesuai dengan deskripsi data, korelasi tiap variabel independen terhadap variabel dependen, serta sebaran data variabel dependen, distribusi data variabel independen.
- k. Mencoba menggunakan beberapa model yang telah ditentukan kemudian melakukan komparasi untuk melihat manakah model terbaik untuk digunakan dalam prediksi.
- l. Mencari model regresi untuk menghasilkan nilai dari variabel dependen.
- m. Membuat deck recommendation slide dan laporan akhir yang di upload di sistem Kampus Merdeka dan GreatEdu.
- n. Setelah menyelesaikan semua serangkaian program pembelajaran Studi Independen, peserta mengikuti sertifikasi BNSP Associate Data Scientist yang telah difasilitasi oleh pihak GreatEdu pada bulan Desember 2022.

II.4 Jadwal Kerja

Program ini berlangsung setiap hari kerja (Senin sampai dengan Jumat) selama 8 jam per harinya, dengan rincian sebagai berikut:

Tabel 2.1 Agenda Kelas

Pukul (WIB)	Durasi (jam)	Aktivitas
-------------	--------------	-----------

08.00 s.d. 11.30	3.5	Kelas Sesi Pagi
13.00 s.d. 16.30	3.5	Kelas Sesi Siang
16.30 s.d. 17.30	1	<i>Self-Study</i>

Program ini berlangsung dari bulan Agustus 2022 sampai dengan bulan Desember 2022. Selama mengikuti rundown kegiatan kelas diselenggarakan secara daring melalui aplikasi *video conference*. dengan rincian sebagai berikut:

Tabel 2.2 Jadwal Kegiatan Pembelajaran

Waktu	Kegiatan	Capaian
Week 1 : 22-27 Agustus	<ul style="list-style-type: none"> • Penyampaian materi Analytical & Critical Thinking • Penyampaian materi Introduction to Data Analytics • Penyampaian materi Problem Solving • Penyampaian materi Coordination & Attention to Detail • Penyampaian materi Data Collection • Penyampaian materi Data Cleaning and Data Preprocessing • Penyampaian materi Dasar Excel untuk Data Analitik • Penyampaian materi Data visualisasi dan dashboard dengan excel dan Tableau 	Memahami konsep dasar data analyst, mengenai pengolahan data dan visualisasi data dengan software komputer. Serta mampu berpikir kritis dan mampu memecahkan masalah dengan kreativitas
Week 2 : 29 Aug - 3 September 2022	<ul style="list-style-type: none"> • Webinar Materi Setting the environment • Webinar Materi Data Wrangling I • Webinar Materi Data Wrangling II • Webinar Materi Practical Statistics I (klasifikasi) 	Memahami data wrangling dan mampu menerapkan dalam data analyst

Week 3 : 5 - 10 September 2022	<ul style="list-style-type: none"> • Webinar Materi Practical Statistics II • Webinar Materi Practical Statistics III • Webinar Materi Must Know Python Tools • Webinar Materi Python Basics I,II,III 	<p>Memahami dan mampu menjalankan python sebagai tools data analyst dan mampu memahami dasar dasar statistik dalam penerapan data analyst</p>
Week 4 12 - 17 September 2022	<ul style="list-style-type: none"> • Webinar Materi Fundamental for coding in python • Webinar Materi NumPy Basics & Panda Basics • Webinar Materi Mathematics for Python • Webinar Materi Panda Basics II 	<p>Memahami dasar coding untuk penggunaan python sebagai software data analyst dan penggunaan matematika dalam python</p>
Week 5 19 - 24 September 2022	<ul style="list-style-type: none"> • Webinar materi Working with Text Files & Working with Text Files II • Webinar materi Working with Text Data • Webinar materi APIs 	<p>Memahami konsep data files dan mampu menggunakan dalam data analyst dan memahami konsep kerja APIs untuk penggunaan data analyst</p>
Week 6 26 September - 1 Oktober 2022	<ul style="list-style-type: none"> • Webinar materi Panda Series • Webinar materi Panda Data frames • Webinar materi Numpy Fundamental • Webinar materi NumPy Data Types 	<p>Memahami dan mampu menjalankan fungsi pandas dan numpy pada pyton untuk pengolahan data</p>
Week 7 3 - 8 Oktober 2022	<ul style="list-style-type: none"> • Webinar materi Working with arrays • Webinar materi Generating data with NumPy • Webinar materi Statistics with NumPy 	<p>Mampu mengolah data untuk mendapatkan mean, median, persentil, std.deviasi, dan korelasi menggunakan python</p>

Week 8 10 - 15 Oktober 2022	<ul style="list-style-type: none"> • Webinar materi NumPy preprocessing • Webinar materi A loan data example with NumPy • Webinar materi A loan data example with NumPy II 	Mampu pengolahan data yaitu input data menggunakan fungsi array, pembersihan data dan editing. dan perhitungan statistik yaitu klasifikasi data dan regresi
Week 9 17 - 22 Oktober 2022	<ul style="list-style-type: none"> • Webinar materi Data visualisation • Webinar materi Data visualisation • Webinar materi Data visualisation • Webinar materi Persiapan Karir 	Bisa menggunakan software untuk visualisasi data seperti google dataset dan mampu menyajikan data dalam bentuk tabel, chart, grafik, diagram, dan lain lain.

Setelah pemaparan materi dan praktek yang di ajarkan oleh expert maupun mentor kegiatan selanjutnya yaitu mengerjakan Project Akhir. dengan rincian sebagai berikut:

Tabel 2.3 Jadwal Pelaksanaan Project Akhir

Tahapan	Aspek yang Dinilai	Start Date	End Date	Durasi (hari)
Business understanding	<ul style="list-style-type: none"> • Memahami bisnis perusahaan yang dijadikan sumber data • Memahami fokus utama dari bisnis tersebut • Menentukan tujuan dan manfaat penelitian 	22 Oktober	25 Oktober	2 hari pertemuan dgn mentor
Data understanding	<ul style="list-style-type: none"> • Memahami informasi 	26 November	1 November	3 hari pertemuan dgn mentor
Data Preparation	<ul style="list-style-type: none"> • Menetapkan attribute yang digunakan 	2 November	10 November	4 hari pertemuan dgn mentor
Modeling	<ul style="list-style-type: none"> • Menetapkan algoritma yang digunakan sesuai dengan tujuan penelitian 	11 November	19 November	4 hari pertemuan dgn mentor

Deployment	<ul style="list-style-type: none">• Taking some action based upon what your model tells you (saran yg bisa diberikan)	20 November	26 Desember	3 hari pertemuan dgn mentor
Making Pitch Deck	<ul style="list-style-type: none">• Desain yang sesuai• Kerapian desain	27 November	3 Desember	3 hari pertemuan dgn mentor

Bab III Project Akhir “Analisis Sentimen Lazada Indonesia Reviews”

III.1 Latar Belakang Proyek Akhir

Lazada didirikan pada tahun 2012, Lazada grup adalah Asia tenggara memimpin Platform eCommerce. Dengan kehadiran di enam negara — Indonesia, Malaysia, Filipina, Singapura, Thailand dan Vietnam – Lazada menghubungkan daerah yang luas dan beragam ini melalui teknologi, logistik dan kemampuan pembayaran. Hari ini, Lazada memiliki pemilihan terbesar Merek dan penjual, dan pada tahun 2030, Lazada bertujuan untuk melayani 300 juta Pelanggan. Pada 2016, Lazada menjadi kapal utama regional Alibaba Grup, dan didukung oleh teknologi terbaik dari Alibaba Infrastruktur.

Tujuan Proyek ini Mengetahui pendapat pengguna Lazada melalui review mereka sebagai bahan evaluasi bagi Lazada (Analisis Sentimen), Mengetahui segmentasi pasar dari pengguna Lazada untuk memberikan saran produk yang lebih sesuai dengan pasar (Klustering), Mengetahui preferensi kegiatan/kategori dari pengguna Lazada sehingga pengguna Lazada lebih mudah menggunakan aplikasi (Klasifikasi).

Berdasarkan latar belakang yang telah diurai diatas, maka dengan ini penulis pembahasan untuk dijadikan topik dalam Laporan final project dengan judul “Analisa Sentimen Lazada Indonesia Reviews”

III.2 Proses Pelaksanaan Proyek Akhir

1	ItemId	category	name	rating	originalRating	reviewTitle	reviewContent	likeCount	upVotes	downVotes	helpful	relevanceScore	boughtDate	clientType	retrievedDate
2	100002528	beli-harddisk-eksternal	Kamal U.,5,,null,bagus mantap dah sesuai pesanan	0,0,0,true	26.51,09 Apr 2019	androidApp	2019-10-02								
3	100002528	beli-harddisk-eksternal	yofanca m.,4,,null,"Bagus, sesuai foto"	0,0,0,true	22.49,24 Sep 2017	androidApp	2019-10-02								
4	100002528	beli-harddisk-eksternal	Lazada Customer,5,,ok mantaapppp barang sesuai pesanan... good,okkkkk mantaaaaaaapppp ... good	0,0,0,true	21.5,04 Apr 2018	androidApp	2019-10-02								
5	100002528	beli-harddisk-eksternal	Lazada Customer,4,,bagus sesuai	0,0,0,true	20.51,22 Sep 2017	androidApp	2019-10-02								
6	100002528	beli-harddisk-eksternal	Yosep M.,5,,null,null	0,0,0,true	16.01,17 Agu 2018	androidApp	2019-10-02								
7	100002528	beli-harddisk-eksternal	Deden,5,,null,null	0,0,0,true	16.01,02 Nov 2017	androidApp	2019-10-02								
8	100002528	beli-harddisk-eksternal	Yeana,5,,null,null	0,0,0,true	13.01,25 Sep 2017	mobile	2019-10-02								
9	100002528	beli-harddisk-eksternal	nurfarida,1,,ada pengirimn ntb bima,bima	4,4,0,true	7.22,	androidApp	2019-10-02								
10	100003785	beli-harddisk-eksternal	Fadjar B.,1,,baru 10 bulan layarnya dah bergaris	0,0,0,true	21.49,06 Apr 2017	androidApp	2019-10-02								
11	100003785	beli-harddisk-eksternal	agung p.,5,,Barang bagus sesuai specs,"Pesan rabu sore,minggu sore sampe,,barang sesuai specs mantab,tkz lazada"	0,0,0,true	19.5,01										
12	100003785	beli-harddisk-eksternal	Hendra W.,5,,null,null	0,0,0,true	18.05 Jan 2018	androidApp	2019-10-02								
13	100004132	beli-harddisk-eksternal	Lazada Guest,1,,Ini cicil pake apa?,"Mau tanya ini cicilnya pake apa ya,cc bkn?"	7,7,0,true	14.88,,	androidApp	2019-10-02								
14	100004132	beli-harddisk-eksternal	Yan M.,5,,Beli cash sesuai Harga di atas,Apakah TV. Tsb. Suda ada anti gores...	1,1,0,true	13.01,,	mobile-app	2019-10-02								
15	100004505	beli-harddisk-eksternal	Andry M.,1,,Pesanan tidak tepat waktu,Pengirim barang tidak sesuai janji. Katanya express 1 hari. Dari hari Jumat barang sampai hari rabu	0,0,0,true	14.99,21 Agu 2018	androidApp	2019-10-02								
16	100004505	beli-harddisk-eksternal	Leonard,4,,TV ok,Kualitas ok	1,1,0,true	22.88,02 Jun 2017	desktop	2019-10-02								
17	100004505	beli-harddisk-eksternal	Lazada Guest,1,,Cara mendapatkan garansi,"Bagaimana tv LED saya merek Sharp 24 hitam LC 24LE170i sekarang mati total baru 2 bln kena	0,0,0,true	22.49,22 Jul 2018	iosApp	2019-10-02								
18	100004505	beli-harddisk-eksternal	Bun D.,3,,null,Pengiriman super lama.. tapi datang juga sikh	0,0,0,true	22.49,22 Jul 2018	iosApp	2019-10-02								
19	100004505	beli-harddisk-eksternal	Fadoli,5,,,"Sharp HD LED TV 24"" - LC-24LE175i - Hitam",ok barang sampai dengan Baik.. smoga terus di pertahankan...	0,0,0,true	21.51,18										
20	100004505	beli-harddisk-eksternal	Akun K.,5,,null,null	0,0,0,true	14.99,21 Agu 2018	androidApp	2019-10-02								
21	100004505	beli-harddisk-eksternal	rina,3,,pengiriman nya laaamaaaaa	null,0,0,0,true	14.99,29 Jul 2018	desktop	2019-10-02								

Gambar 3.1 Dataset Reviews Lazada Indonesia

itemid	category	name	brandName	url	price	averageRating	totalReviews	retrievedDate
100002528	beli-harddisk-eksternal	TOSHIBA Smart HD LED TV 32"	Hitam	https://www.lazada.co.id/products/toshiba-smart-hd-led-tv-32-32l5650vj-free-				
100003785	beli-harddisk-eksternal	TOSHIBA Full HD Smart LED TV 40"	Hitam	https://www.lazada.co.id/products/toshiba-full-hd-smart-led-tv-40-40l5650vj-hitam-khusu-				
100004132	beli-harddisk-eksternal	Samsung 40 Inch Full HD Flat LED Digital TV 40J5000	LG	https://www.lazada.co.id/products/samsung-40-inch-full-hd-flat-led-digital-tv-40j5000-i100004132-s100006644.html?sear-				
100004505	beli-harddisk-eksternal	Sharp HD LED TV 24"	Hitam	https://www.lazada.co.id/products/sharp-hd-led-tv-24-lc-24le175i-hitam-i100004505-s100007387.html?search=1,1275000,3,11,2-				
100005037	beli-harddisk-eksternal	Lenovo Ideapad 130-15AST LAPTOP MULTIMEDIA	AMD A4-9125 8GB DDR4 1TB HDD AMD Radeon R3	https://www.lazada.co.id/products/lenovo-ideapad-130-15ast-laptop-multimedia-i-amd-a4-9125-8gb-ddr4-1tb-hdd-amd-radeon-r3-i15,6-hd-led-i-dvdrw-i-dos-i100005037-lenovo-i100006686-				
100006868	beli-harddisk-eksternal	Sandisk Flashdisk Cruzer Glide CZ60 16GB	iRing Mobile Phone	https://www.lazada.co.id/products/sandisk-flashdisk-cruzer-glide-cz60-16gb-iring-mobile-phone-i100006868-				
100007498	beli-harddisk-eksternal	Asus X407UB-BV187T Laptop Multimedia Murah	Core i5-8250U 4GB DDR4 1TB HDD + SLOTT SSD M2	https://www.lazada.co.id/products/asus-x407ub-bv187t-laptop-multimedia-murah-i-core-i5-8250u-4gb-ddr4-1tb-hdd-sloTT-ssd-m2-i-nvidia-geforce-mx110-2gb-i-windows-10-i-no-dvdrw-i-14.0-hd-uslim-i-asus-i100007498-asus-i100006061-				
100010722	beli-harddisk-eksternal	Philips 32PHA3052S/70 32" Televisi LED (FREE Bracket LED)	Hitam	https://www.lazada.co.id/products/philips-32pha3052s70-32-televiSi-led-free-bracket-led-i100010722-philips-i100006061-				
100010745	beli-harddisk-eksternal	Philips 39PHA4251S/70 39" Televisi LED (FREE Bracket LED)	Hitam	https://www.lazada.co.id/products/philips-39pha4251s70-39-televiSi-led-free-bracket-led-i100010745-philips-i100006061-				
100015720	beli-harddisk-eksternal	Acer E5 575G-74E2 - Core i7-6500U - RAM 8GB DDR4 - HDD 1TB - VGA Nvidia 940MX 2GB - Black	Acer	https://www.lazada.co.id/products/acer-e5-575g-74e2-core-i7-6500u-ram-8gb-d-d-100015720-acer-i100006061-				
100024235	beli-harddisk-eksternal	Sharp LED 32 LE 180 - 32 Inc - Hitam KHUSUS JABODETABEK (untuk keluar kota wajib di packing kayu)	Sharp	https://www.lazada.co.id/products/sharp-led-32-le-180-32-inc-hitam-kh-k-100024235-sharp-i100006061-				
100027329	beli-harddisk-eksternal	SanDisk Flashdisk Cruzer Glide CZ60 - 32GB - Hitam Merah - USB 2.0	SanDisk	https://www.lazada.co.id/products/sandisk-flashdisk-cruzer-glide-cz60-32gb-hitam-merah-usb-20-i100027329-sandisk-i100006061-				
100027616	beli-harddisk-eksternal	Maxtor Hardisk Eksternal M3 1TB 2.5" USB 3.0 HDD EXT Storage	Maxtor	https://www.lazada.co.id/products/maxtor-hardisk-eksternal-m3-1tb-25-usb-30-hdd-ext-storage-i100027616-maxtor-i100006061-				
100028438	beli-harddisk-eksternal	SanDisk Flashdisk Ultra CZ48 100MB/s - 32GB - Merah - USB 3.0	SanDisk	https://www.lazada.co.id/products/sandisk-flashdisk-ultra-cz48-100mbs-32gb-merah-usb-30-i100028438-sandisk-i100006061-				
100028438	beli-harddisk-eksternal	Orico PHL-25 2.5 Inch HDD Protection Case Bag - Hitam	Orico	https://www.lazada.co.id/products/orico-phl-25-25-inch-hdd-protection-case-bag-hitam-i100028438-orico-i100004187-ht-				
100030605	beli-harddisk-eksternal	SanDisk Flashdisk Cruzer Switth CZ52 - Flashdisk 64GB - Hitam Merah - USB 2.0	SanDisk	https://www.lazada.co.id/products/sandisk-flashdisk-cruzer-switch-cz52-flashdisk-64gb-hitam-merah-usb-20-i100030605-sandisk-i100006061-				
100034104	beli-harddisk-eksternal	SanDisk Flashdisk 64 GB Cruzer Blade CZ50 - Hitam	SanDisk	https://www.lazada.co.id/products/sandisk-flashdisk-64-gb-cruzer-blade-cz50-hitam-i100034104-s100048518.html?sear-				
100034117	beli-harddisk-eksternal	SanDisk Flashdisk 64GB Cruzer Blade CZ50 - Hitam + iRing Mobile Phone warna Random	SanDisk	https://www.lazada.co.id/products/sandisk-flashdisk-64gb-cruzer-blade-cz50-hitam-iring-mobile-phone-warna-random-i100034117-sandisk-i100006061-				
100036366	beli-harddisk-eksternal	Sharp HD LED TV 24"	Hitam	https://www.lazada.co.id/products/sharp-hd-led-tv-24-lc-24le170i-hitam-i100036366-s100051598.html?search=1,1149000,4,54,2-				
100038665	beli-harddisk-eksternal	Lenovo Ideapad 110 - 15IBR - Windows 10 - Celeron N3060 - Ram 4GB - 500GB - 15,6"	Hitam	https://www.lazada.co.id/products/lenovo-ideapad-110-15ibr-windows-10-celeron-n3060-ram-4gb-500gb-156-i100038665-lenovo-i100006061-				

Gambar 3.2 Dataset Items Lazada Indonesia

- Data yang digunakan yaitu data Reviews dan Item dari Lazada di Indonesia. Data Reviews berjumlah 203788 dan 15 variabel yang terdiri data
 itemid,category,name,brandName,url,price,averageRating,totalReviews,retrievedDate. Kalau Data Items berjumlah 10947 dan 9 variabel yang terdiri data
 itemId,category,name,brandName,url,price,averageRating,totalReviews,ret
 rievedDate
- Dalam proses pengerjaan final proyek, kami menerapkan metode CRISP DM yang dimana dimulai dari:
 - Bussiness understanding*: dalam *bussiness undestanding* kami memahami bisnis dari lazada itu sendiri sehingga nantinya dapat dengan mudah untuk melakukan proses analisis
 - Data understanding*: pada tahapan ini kami memahami setiap variabel yang ada pada dataset. Pada dataset yang digunakan terdapat 15 variabel pada dataset reviews dan 9 variabel pada dataset items.
 - Data preparation*, pada tahapan ini yang dilakukan adalah Ubah tipe data:
 - Kami mengubah jenis tanggal pengambilan dan pemutakaan dari objek ke waktu kencana
 - Pilih yang berguna untuk analisa ini
 - Kami menggunakan penilaian dan reviewContent untuk analisis keluhan
 - Hapus stopword
Kami menghapus kata yang dapat dikatakan tidak penting.. Contoh: 'yang', 'untuk', 'nya', ' ', ' ', 'lazada' DLL
 - Filtering
Dalam langkah ini, kita menghapus tanda baca, URL, gambar dan membuat kata pada kasus yang lebih rendah
 - Stemming

- Proses pengurangan kata ke batang katanya yang awalan ke akhiran dan awalan atau ke akar kata yang dikenal sebagai lemma
- label

Label berdasarkan penilaian yang diberikan oleh pelanggan. Jika memiliki 5 atau 4 peringkat dapat dikatakan itu tinjauan positif dan nilai 1. Jika memiliki peringkat 3 atau di bawah dapat dikatakan itu tinjauan negatif dan nilai 0.
 - Split Data

Membagi data menjadi data pelatihan dan pengujian. Perbandingan yang digunakan adalah 80% : 20%
 - penanaman kata

Metode yang digunakan untuk membenamkan kata adalah TF-IDF (dokumen indeks frekuensi istilah Smote (sintetis minoritas Oversampling teknik) Ini digunakan untuk menangani data ketidakseimbangan.
 - Klasifikasi
 - evaluasi
- d. *Modelling*: pada tahapan inilah analisis dimulai. Di sini metode yang digunakan yaitu *Analysis Sentimen* dan untuk toolsnya menggunakan Google Colab, Tahapan-tahapan dalam analisis yang dilakukn sebagai berikut:

- Mengimport *library* yang akan digunakan

```
[ ] import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.ticker as tick
import seaborn as sns
import re
import os

%matplotlib inline

sns.set_style('white')
```

Gambar 3.3. Mengimport library di Google Colab

- Memanggil dataset

```
[ ] from google.colab import drive
drive.mount('/content/drive')

[ ] df_items = pd.read_csv('/content/drive/MyDrive/Final SI/20191002-items.csv')
df_reviews = pd.read_csv('/content/drive/MyDrive/Final SI/20191002-reviews.csv')
```

Gambar 1.4 Input dataset yang digunakan

- Mengecek *missing value* dari dataset

```
[ ] df_reviews.isnull().sum()
itemId          0
category        0
name            0
rating          0
originalRating  35533
reviewTitle     27804
reviewContent   16498
likeCount       0
upVotes         0
downVotes       0
helpful         0
relevanceScore  0
boughtDate      1623
clientType      0
retrievedDate   0
dtype: int64
```

Gambar 3.5. Mengecek missing value

- Melihat tipe data

```
[ ] df_reviews.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 35535 entries, 0 to 35534
Data columns (total 15 columns):
#   Column          Non-Null Count  Dtype
---  -
0   itemId          35535 non-null   int64
1   category        35535 non-null   object
2   name            35535 non-null   object
3   rating          35535 non-null   int64
4   originalRating  2 non-null      float64
5   reviewTitle     7731 non-null   object
6   reviewContent   19037 non-null  object
7   likeCount       35535 non-null   int64
8   upVotes         35535 non-null   int64
9   downVotes       35535 non-null   int64
10  helpful         35535 non-null   bool
11  relevanceScore  35535 non-null   float64
12  boughtDate      33912 non-null  object
13  clientType      35535 non-null  object
14  retrievedDate   35535 non-null  object
dtypes: bool(1), float64(2), int64(5), object(7)
memory usage: 3.8+ MB
```

Gambar 3.6. Mengecek missing value

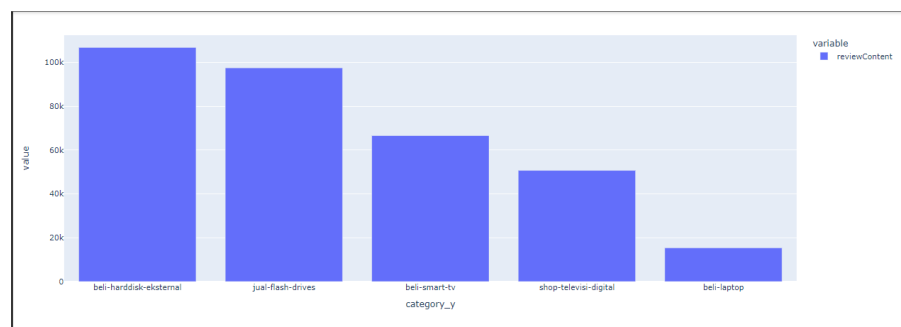
- Melakukan eksplorasi data untuk mengetahui info apa saja yang ada pada data.

Melihat category dengan review paling banyak

```
[ ] #melihat category dengan review paling banyak
project.groupby(["category_y"])["reviewContent"].count().sort_values(ascending=False)

category_y
beli-harddisk-eksternal    106871
jual-flash-drives         97548
beli-smart-tv              66610
shop-televisi-digital     50718
beli-laptop                15350
Name: reviewContent, dtype: int64
```

Gambar 3.7. review paling banyak

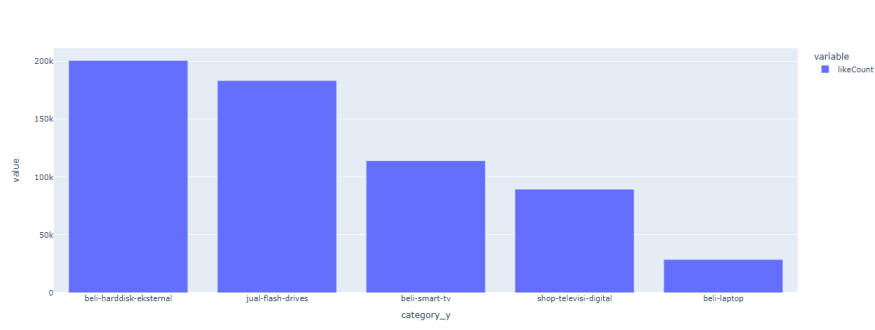


Gambar 3.8. diagram review paling banyak

```
[ ] #melihat category dengan like paling banyak
project.groupby(["category_y"])["likeCount"].count().sort_values(ascending=False)

category_y
beli-harddisk-eksternal    200731
jual-flash-drives         183364
beli-smart-tv              114021
shop-televisi-digital     89384
beli-laptop                28600
Name: likeCount, dtype: int64
```

Gambar 3.9. melihat category dengan like paling banyak

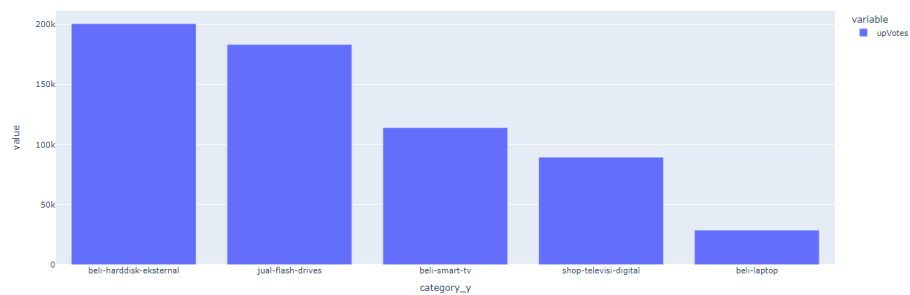


Gambar 3.10. diagram category dengan like paling banyak

```
[ ] #melihat category dengan upvote paling banyak
project.groupby(["category_y"])["upVotes"].count().sort_values(ascending=False)

category_y
beli-harddisk-eksternal    200731
jual-flash-drives         183364
beli-smart-tv              114021
shop-televisi-digital     89384
beli-laptop                28600
Name: upVotes, dtype: int64
```

Gambar 3.11. melihat category dengan upvote paling banyak



Gambar 3.12. melihat category dengan upvote paling banyak

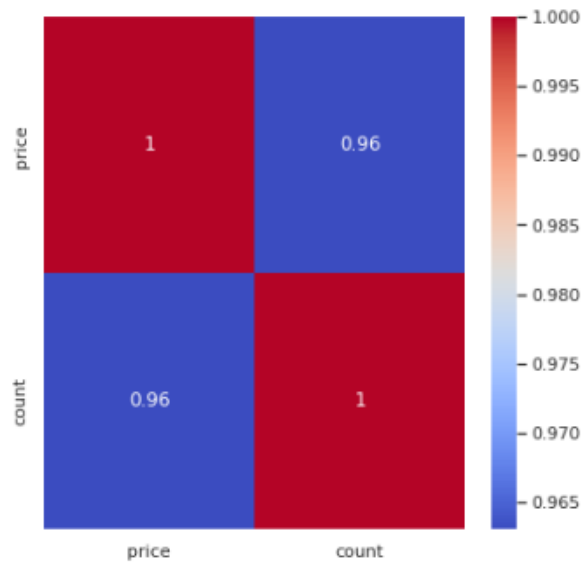
- Visualisasi Piecat



Gambar 3.13. Komposisi items menurut jumlah kategorinya

- Melihat korelasi dari setiap variabel

Tabel Korelasi Jumlah Transaksi Harian dan Nilai Transaksinya



Gambar 3.14. Korelasi jumlah transaksi harian dan nilai transaksi

Berdasarkan visualisasi resampling tersebut, pergerakan data antara jumlah transaksi dan nilai transaksi hampir sama. Keduanya sama-sama naik dengan sangat signifikan menjelang tahun 2019. Ketika kita melakukan visualisasi secara mendetail pada sampel periode Juli-September 2018, keduanya sama-sama memiliki tren yang menyerupai. Selanjutnya kita mencoba korelasikan kolom count dan price, dan menghasilkan korelasi sebesar 0.96 yang berarti memiliki korelasi positif yang sangat kuat.

- WordCloud

Bab IV Penutup

IV.1 Kesimpulan

Berdasarkan dari implementasi kegiatan dalam mengikuti Studi Independen Bersertifikat Kampus Merdeka. Dari ini dapat disimpulkan sebagai berikut:

- a. Pelaksanaan kegiatan Studi Independen Bersertifikat Kampus Merdeka dari GreatEdu berjalan dengan lancar dan baik selama kurang lebih 5 bulan dari 18 Agustus 2022 hingga 31 Desember 2022.
- b. Materi yang di berikan sudah sangat baik dan sudah sesuai dengan kurikulum yang diberikan, serta sesuai dengan bidang konsentrasi yang saya ambil
- c. Penjelasan materi terkait Data Analyst ini dijelaskan oleh expert yang berpengalaman di bidangnya
- d. Dalam masa pembelajaran, seluruh peserta akan diberikan mentor yang dibagi ke dalam 5 kelompok untuk review tugas mingguan dan materi yang telah dijelaskan oleh expert. Kemudian dari setiap kelompok mentoring akan dibagi lagi menjadi 5 kelompok yang berisi 4-5 peserta untuk pengerjaan project akhir
- e. Project akhir yang dilakukan oleh kelompok penulis yaitu analisa sentimen lazada indonesia reviews yang ditujukan untuk mengetahui reviews positif dan negatif dari pelanggan lalu untuk mengetahui barang mana saja yang paling sering dibeli.

IV.2 Saran

Adapun saran mengenai proses pelaksanaan MSIB PT Greatedu Global Mahardika, yaitu :

- a. PT Greatedu Global Mahardika harus tetap mempertahankan dan meningkatkan kedisiplinan yang telah ada

- b. Senantiasa menjalin hubungan baik dengan berbagai instansi, lembaga, maupun perusahaan yang berpotensi mengembangkan pengetahuan dan wawasan peserta yang akan melaksanakan program MSIB
- c. Tetap ikut serta sebagai mitra dalam program MSIB batch selanjutnya
- d. Kedepan tetap terjalin hubungan antara mentor, teman beserta seluruh karyawan dari PT Greatedu Global Mahardika
- e. Waktu yang diberikan untuk pengerjaan final project terlalu singkat dan sebaiknya untuk kedepannya bisa ditambahkan lagi waktu pengerjaannya agar hasilnya bisa lebih maksimal
- f. Sebaiknya untuk jadwal bimbingan final project terkhusus bagian teknis lebih diperhatikan lagi agar proses pengerjaan final project lebih baik.

Bab V Referensi

- [1] Ghazal, T.M., Hussain, M.Z., Said, R.A., Nadeem, A., Hasan, M.K., Ahmad, M., Khan, M.A. and Naseem, M.T., 2021. Performances of K-means clustering algorithm with different distance metrics.
- [2] Raiter, O., 2021. Segmentation of Bank Consumers for Artificial Intelligence Marketing. *International Journal of Contemporary Financial Issues*, 1(1), pp.39-54.
- [3] Anggarwati, D., Nurdiawan, O., Ali, I. and Kurnia, D.A., 2021. Penerapan Algoritma K-Means Dalam Prediksi Penjualan Karoseri. *JURNAL DATA SCIENCE & INFORMATIKA*, 1(2), pp.58-62.
- [4] Azis, N., Wahidin, A.J., Cakranegara, P.A., Muditomo, A. and Efendi, E., 2022. Visualization Of Tourist Visit Time Series Data Using Google Data Studio. *Jurnal Mantik*, 6(2), pp.2153-2159.
- [5] Huber, S., Wiemer, H., Schneider, D. and Ihlenfeldt, S., 2019. DMME: Data mining methodology for engineering applications—a holistic extension to the CRISP-DM model. *Procedia Cirp*, 79, pp.403-408.

Bab VI Lampiran A. TOR

TERM OF REFERENCE (TOR) STUDI INDEPENDEN BERSERTIFIKAT DATA ANALYTICS DI GREATEDU

A. Tema Program

Ilmu Data

B. Latar Belakang Program

GreatEdu merupakan perusahaan yang bergerak di bidang teknologi, lebih tepatnya education technology. Sebagai perusahaan yang bergerak di bidang digital, GreatEdu menyadari pentingnya data untuk terus berkembang. Data yang dapat menjadi acuan dalam pengambilan keputusan oleh pimpinan perusahaan. Bagaimana data tersebut bisa diolah, divisualisasi, dan dimanfaatkan dengan baik.

Profesi Data analis merupakan profesi dengan permintaan yang cukup banyak pada tahun ini hingga beberapa tahun ke depan. Mengutip dari Data Careers Asia, ada 170,000 lowongan sebagai Data Analis di USA. Sedangkan, menurut Glints, di Indonesia sendiri ada 9,428 lowongan Data Analis di bulan Oktober 2022. Selain lowongan yang terbuka lebar, gaji seorang Data Analis pun sangat menjanjikan, dikutip dari www.salaryexplorer.com, rata-rata gaji per bulan dari Data Analis mencapai Rp 11,000,000. Sayangnya, dikarenakan belum adanya sekolah khusus data, maka perlu ada pelatihan bagi seorang calon data analis untuk meningkatkan keterampilan yang dibutuhkan. Jikalau ada, hanya basicnya saja dan itu tersebar di beberapa jurusan. Tidak ada ketentuan khusus bagi seseorang yang mau membangun karir sebagai data analyst, semua orang dengan latar belakang apapun mampu, asal pembelajaran selama pelatihan ditekuni dengan baik.

Namun, berdasarkan pengalaman expert yang sudah lama berkecimpung pada bidang Data Analisis, masalah ini justru menjadi peluang bagi siapapun yang tertarik pada bidang ini. Adapun, skill gap yang terjadi pada seorang analyst adalah belum mampu mengimplementasikan metode analisis sebagai poin wajib yang mesti dikuasai oleh seorang data analyst. Oleh karena itu, GreatEdu menginisiasi tema data pada program SIB Kampus Merdeka.

C. Visi dan Misi

Menjadi sarana dalam membentuk ekosistem pendidikan masa depan khususnya di bidang teknologi.

D. Peserta

Peserta program ini adalah mahasiswa yang berasal dari Perguruan Tinggi di bawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.

E. Model dan Metode Belajar

Synchronous dan Asynchronous. Metode belajar synchronous dengan Zoom atau Google Meet dengan metode ceramah, demonstrasi, simulasi, dan atau studi kasus, sedangkan metode belajar asynchronous dengan membaca modul, menonton video pembelajaran, mengerjakan kuis, dan praktik mandiri. Tujuan dari program ini adalah peserta mampu membuat deck rekomendasi/ insight dari hasil analisis dan pengolahan data.

F. Learning Path

1. **Agustus - Oktober**, pemberian materi oleh Expert melalui metode pembelajaran synchronous dan asynchronous. Selain itu ada sesi review bersama Mentor dan sesi refleksi.
2. **November - Desember**, mengerjakan final project berupa deck rekomendasi sesuai dengan sumber data yang diberikan.

G. Tujuan Program

Tujuan dari program ini adalah peserta mampu membuat deck rekomendasi/ insight dari hasil analisis dan pengolahan data.

Partnership Officer,



Robert Edy Sudarwan

Bekasi, 28 Desember 2022
Peserta Program,



M. Anwar Lirboyo
201910225319

Bab VII Lampiran B. Log Activity

Minggu/Tgl	Kegiatan	Hasil
Ke 1 (18-19 Agustus 2022)	<ul style="list-style-type: none"> • Onboarding pisah sambut angkatan 2 dan 3 program MSIB • Pembuatan akun di website GreatEdu dan belajar mandiri pada modul yang telah diberikan 	<ul style="list-style-type: none"> • Mempelajari dan mengetahui mengenai Data Mining for the Masses melalui modul yang telah disediakan pada website GreatEdu
Ke 2 (22-26 Agustus 2022)	<ul style="list-style-type: none"> • Webinar Analytical and Critical Thinking • Webinar Introduction to Data Analytics • Webinar Problem Solving • Webinar coordination Attentiion to detail • Webinar data collection and preprocessing • Webinar Dasar Excel untuk data analytic dan visualisasi data menggunakan excel dan dashboard pada tableau • Review tugas praktek mingguan 	<ul style="list-style-type: none"> • Mengetahui keterampilan penting yang harus dimiliki oleh seorang data analytic yaitu berpikir kritis dan mempelajari bagaimana kiat-kiat dalam berpikir kritis • Mengetahui apa saja teknik dan juga tools yang digunakan dalam analisis data • Sebagai seorang data analytic harus memiliki pemahaman yang baik mengenai problem solving • Memahami bahwa pentingnya koordinasi sesama team dalam pekerjaan • Mengetahui bagaimana langkah-langkah dalam mengumpulkan data dan mempersiapkan data

		<ul style="list-style-type: none"> • Mengetahui dasar excel yang biasa digunakan data analytic • Mengetahui cara visualisasi yang baik dan menarik menggunakan excel dan tablu
Ke 3 (29 Agustus-3 September 2022)	<ul style="list-style-type: none"> • Webinar materi setting the environment • Webinar materi Data wrangling I • Webinar Data Wrangling II • Webinar Materi Practical Statistics I • Review tugas praktek mingguan 	<ul style="list-style-type: none"> • Melakukan instal anaconda dan setting anaconda • Mengetahui cara memperoleh data dan membaca data csv • Mengetahui bagaimana cara mengoperasikan pandas data frame • Mengetahui basic SQL, bagaimana cara cleansing data dan cara melakukan EDA yang benar dan menarik • Memahami dasar-dasar statistik
Ke 4 (5-10 September 2022)	<ul style="list-style-type: none"> • Webinar Materi Practical Statistics II • Webinar Materi Practical Statistics III • Webinar Materi Must Know Python Tools • Webinar Materi Python basic I • Webinar Materi Python basic II • Webinar Materi Python basic 	<ul style="list-style-type: none"> • Melanjutkan pembeajaran mengenai dasar-dasar statistik yang harus dikuasai oleh seorang data analytic • memahami metode analisis yang digunakan oleh data analytic • Mengetahui apa itu python • Menegtahui dasar-dasar python yang harus dikuasi

	<p>III</p> <ul style="list-style-type: none"> • Review tugas praktek mingguan 	<p>oleh data analytic</p>
Ke 5 (12 -17 September)	<ul style="list-style-type: none"> • Webinar Materi Fundamental for coding in python • Webinar Materi NumPy Basics & Panda Basics • Webinar Materi Mathematics for Python • Webinar Materi Panda Basics II • Review tugas praktek mingguan 	<ul style="list-style-type: none"> • Memahami module, package dan library pada python yang biasa digunakan oleh data analytic • Memahami dokumentasi pada python • Mempelajari pemrograman yang berorientasi pada objek • Mengetahui apa itu pandas • Mempelajari cara membuat data frame, dropping, sort dan rank, perbedaan parameter dan argument pada pandas • Memahami kenapa menggunakan numpy dan bagaimana cara menggunakannya di python
Ke 6 (19-24 September 2022)	<ul style="list-style-type: none"> • Webinar materi Working with Text Files & Working with Text Files II • Webinar materi Working with Text Data • Webinar materi APIs • Review tugas praktek mingguan 	<ul style="list-style-type: none"> • Mengetahui bagaimaa cara bekerja dengan file di python • Mengetahui perbedaan file dan objek • Mengetahui perbedaan structur, semi structur dan unstruktur data • Menegtahui berbagai cara import file di python

<p>Ke 7 (26 September–1 Oktober 2022)</p>	<ul style="list-style-type: none"> • Webinar materi Panda Series • Webinar materi Panda Data frames • Webinar materi NumPy Fundamental • Webinar materi NumPy Data Types • Review tugas praktek mingguan 	<ul style="list-style-type: none"> • Memahami dan mempelajari bagaimana cara membuat series • Mengetahui function apa saja yang bisa diaplikasikan di series • Mempelajari bagaimana cara menggunakan method dan function secara berurutan • Mengetahui attribute dalam dataframe • Mengetahui agregasi dalam dataframe • Mengetahui bagaimana cara memfilter dataframe • Mengetahui bagaimana cara indexing di numpy • Mengetahui tipe-tipe data di numpy • Mengetahui bagaimana cara mengecek tipe data array • Mengetahui perbedaan array dan list
<p>Ke 8 (3-8 Oktober 2022)</p>	<ul style="list-style-type: none"> • Webinar materi Working with arrays • Webinar materi Generating data with NumPy • Webinar materi Statistics with NumPy • Review tugas praktek 	<ul style="list-style-type: none"> • Mempelajari bagaimana cara slicing pada Numpy • Mengetahui bagaimana cara memanggil data random dengan numpy • Mempelajari bagaimana penerapan numpy pada basic

	mingguan	statistic seperti cara menampilkan nilai max, min, rata-rata, varians, korelasi dan histogram
Ke 9 (10-15 Oktober	<ul style="list-style-type: none"> • Webinar materi NumPy preprocessing • Webinar materi A loan data example with NumPy • Webinar materi A loan data example with NumPy II • Review tugas praktek mingguan 	<ul style="list-style-type: none"> • Mempelajari bagaimana memproses data dengan numpy seperti membentuk array, menggabungkan array, menemukan sebuah array dan mengurutkan array • Mempelajari bagaimana cara menemukan, mengganti, dan melakukan cleaning data dengan numpy
Ke 10 (17-22 Oktober 2022)	<ul style="list-style-type: none"> • Webinar materi Data visualisation I • Webinar materi Data visualisation II • Webinar materi Data visualisation III • Webinar materi Persiapan Karir • Review tugas praktek mingguan 	<ul style="list-style-type: none"> • Mempelajari bagaimana cara menyajikan hasil visualisasi data • Mempelajari bagaimana cara menggunakan google data studio • Mempelajari bagaimana cara visualisasi dasar dengan dashboard google data studio • Memahami bagaimana cara menggunakan dimensi dan metriks pada google data studio • Mempelajari bagaimana cara menggunakan filter dasar pada google data studio

		<ul style="list-style-type: none"> • Mengenal apa itu halaman, advance filters dan clickable pada looker data studio • Mempelajari bagaimana cara mengembangkan bentuk visualisasi pada looker studio • Mempelajari bagaimana cara membuat dashboard yang interaktif dan efektif
Ke 11 (24-29 Oktober 2022)	<ul style="list-style-type: none"> • Final Project 	<ul style="list-style-type: none"> • Melakukan tahapan bussiness understanding yaitu memahami bisnis perusahaan yang dijadikan sumber data • Memahami fokus utama dari bisnis perusahaan • Menentukan tujuan dan manfaat
Ke 12 (31 Oktober – 5 November 2022)	<ul style="list-style-type: none"> • Final Project 	<ul style="list-style-type: none"> • Melakukan tahapan data understanding yaitu memahami informasi pada data • Melakukan proses EDA (Exploratory Data Analysis)
Ke 13 (7-12 November 2022)	<ul style="list-style-type: none"> • Final Project 	<ul style="list-style-type: none"> • Melakukan tahapan data preparation yaitu menentukan atribut apa saja yang akan digunakan pada proses analisis
Ke 14 (14-19 November)	<ul style="list-style-type: none"> • Final Project 	<ul style="list-style-type: none"> • Melakukan tahapan modelling yaitu menentukan

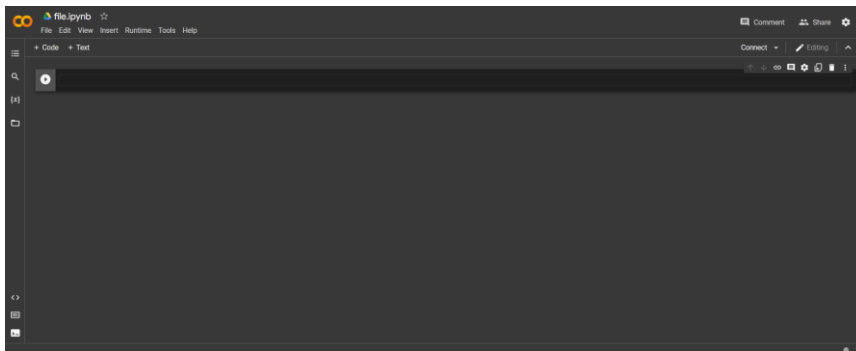
2022)		<p>algoritma apa yang akan digunakan pada penelitian</p> <ul style="list-style-type: none"> • Melakukan proses analisis data dengan algoritma yang telah ditentukan
Ke 15 (21-26 November 2022)	<ul style="list-style-type: none"> • Final Project 	<ul style="list-style-type: none"> • Melakukan tahapan deployment dimana memberikan saran berdasarkan hasil modelling atau hasil analisis yang telah dilakukan
Ke 16 (28 November – 3 Desember 2022)	<ul style="list-style-type: none"> • Final Project 	<ul style="list-style-type: none"> • Menentukan desain ppt yang interaktif dan menarik • Persiapan untuk presentasi hasil penelitian

Bab VIII Lampiran C. Dokumen Teknik

a. Beberapa tampilan software yang digunakan:

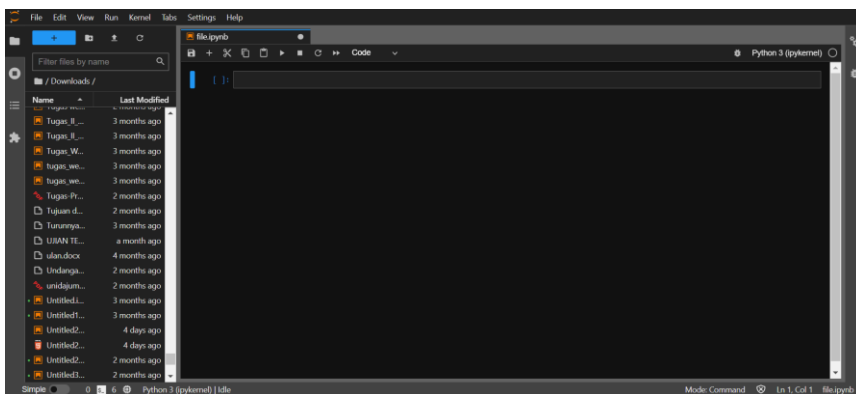
1. Colaboratory

Google Colaboratory adalah executable document yang memungkinkan kamu dalam menulis, mengedit, serta membagikan program yang sudah disimpan pada drive maupun yang baru.



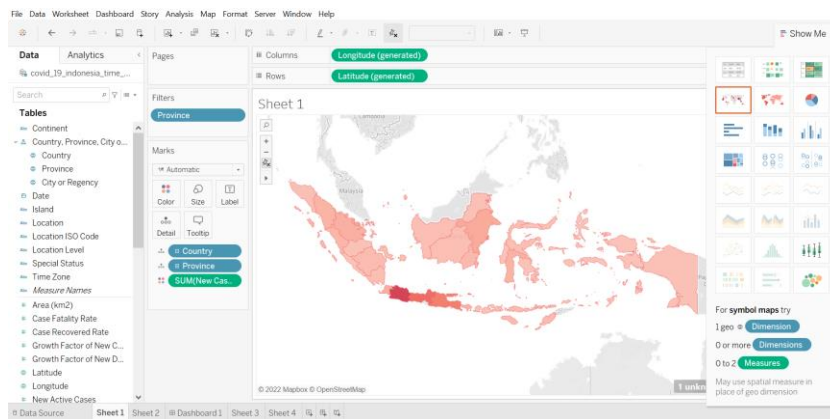
2. Jupyter lab

Jupyter lab adalah aplikasi web sumber terbuka yang memungkinkan membuat dan berbagi dokumen yang berisi kode langsung, persamaan, visualisasi, dan teks naratif. Penggunaan termasuk pembersihan dan transformasi data, simulasi numerik, pemodelan statistik, visualisasi data, dan pembelajaran mesin.



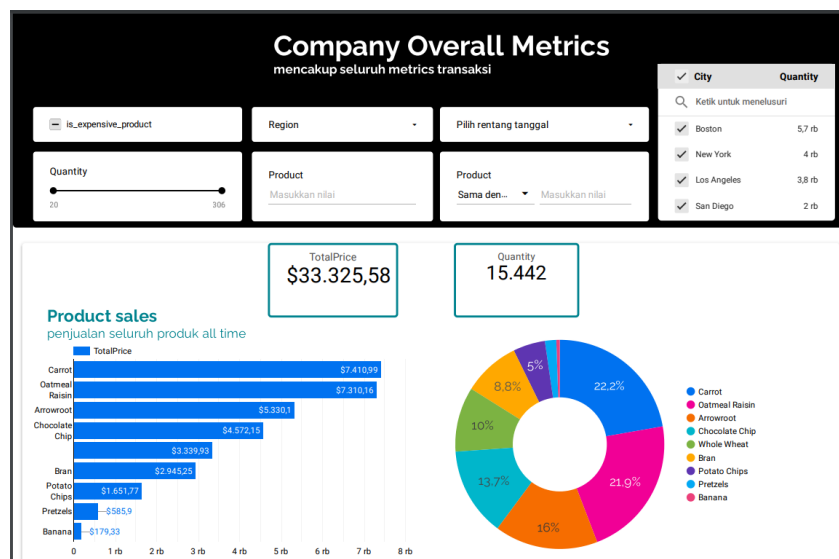
3. Tableau

Tableau adalah perangkat lunak yang dipakai oleh perusahaan untuk memproses data-data yang ada. Platform ini memungkinkan adanya proses kolaborasi kreatif dari berbagai divisi perusahaan dalam rangka menampilkan visualisasi data yang menarik untuk dilihat dan mudah untuk dipahami.



4. Google data studio-Looker Studio

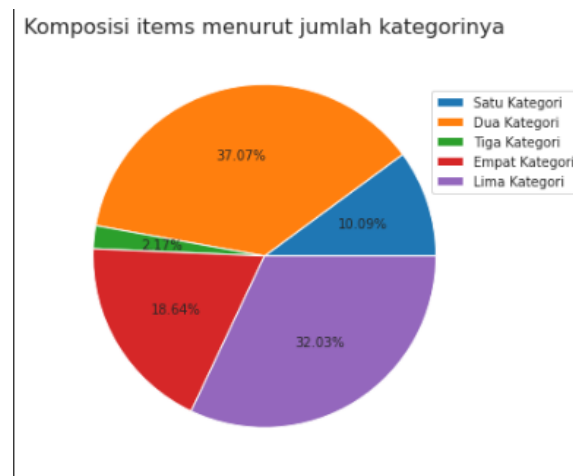
Google data studio-Looker Studio adalah alat gratis untuk mengumpulkan dan membuat visualisasi data yang menarik, dapat disesuaikan, interaktif, dan responsif. Kamu dapat membuat panel dari template, menggunakan grafik, peta, dan tabel yang paling beragam.



- b. Hasil final project
1) Judul final projek

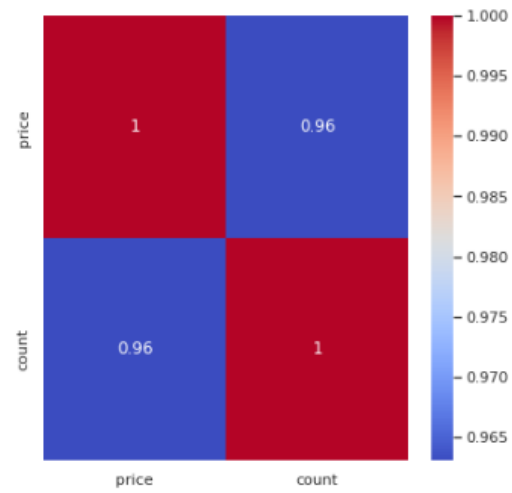


- 2) Visualisasi data

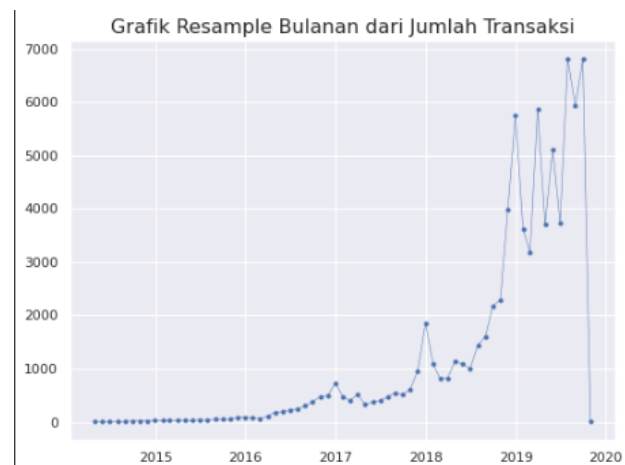


- 3) Korelasi jumlah transaksi harian dan nilai transaksi

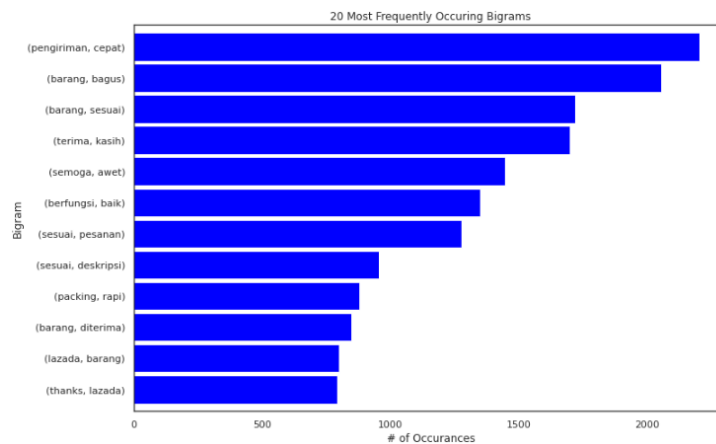
Tabel Korelasi Jumlah Transaksi Harian dan Nilai Transaksinya



4) Grafik resample bulanan dari jumlah transaksi



5) 20 bigram yang paling sering muncul



C. Lampiran Tugas Selama Kegiatan

1. Assignment I – Introduction to Data Analytics

NAMA : MUHAMMAD ANWAR LIRBOYO

NPM : 201910225319

KELAS : DATA ANALYTICS

- Jawab pertanyaan di bawah dengan jujur. Silahkan tulis jawaban melalui google document atau word, kemudian attach link ataupun file. Good luck!
 1. Apa yang kamu ketahui tentang Jobdesk utama untuk seorang Data Analyst?
 2. Tools apa saja yang biasa digunakan untuk seorang Data Analyst?
 3. Apa perbedaan antara Data Scientist dengan Data Analyst?

Jawaban

1. Adalah seseorang yang bertanggung jawab untuk melakukan analisis dan riset data (data analytics) dengan menggunakan tools tertentu.
2. – Apache Park
 - Datapine
 - SQL
 - SAS
 - EXCEL
 - TABLEAU

2. Assignment II – Analytical Thinking

NAMA : MUHAMMAD ANWAR LIRBOYO

NPM : 201910225319

KELAS : DATA ANALYTICS

ANALYTICAL THINKING

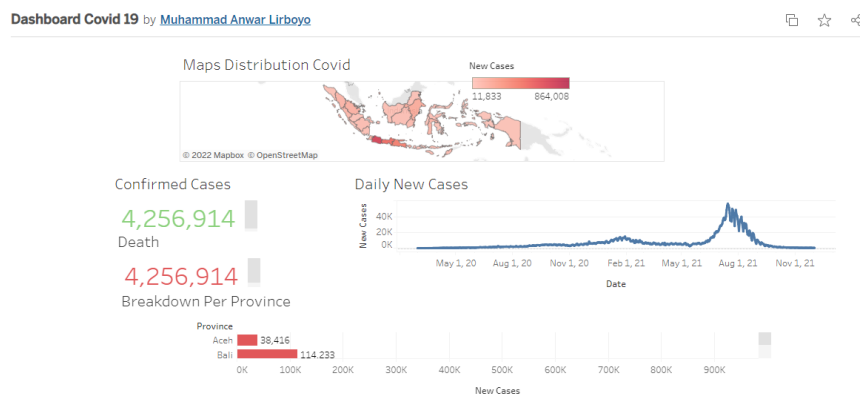
Isilah pertanyaan dengan jujur dan cermat. Pertanyaan dapat dilihat di dokumen PDF pada tautan di bawah ini. Silahkan tulis jawaban melalui google document atau word, kemudian attach link ataupun file. Good luck!

1. Anda sedang makan siang dengan seorang teman lama dari universitas, dan dia mencari beberapa saran bisnis. Dia sedang berpikir untuk membuka kedai kopi di area Jabodetabek.

Dia melihat potensi dalam bisnis ini tetapi menginginkan bantuan anda dalam menentukan apakah membuka kedai kopi adalah ide yang bagus.

Dari kasus di atas,

4. Assignment III - Data Visualization with Tableau



5. Assignment IV – Problem Solving

Tugas problem solving yang di upload ke classroom

NAMA : MUHAMMAD ANWAR LIRBOYO

NPM : 201910225319

KELAS : DATA ANALYTICS

TUGAS PROBLEM SOLVING

1. Netflix, suatu perusahaan video on demand mengalami penurunan jumlah pengguna dalam beberapa bulan terakhir.
 - **Bagaimana rekomendasi anda agar produk tersebut kembali mengalami kenaikan dari sisi jumlah pengguna?**
Promosi secara berkala secara menarik agar banyak pelanggan berdatangan kembali dan diberi diskon agar penawaran yang diberikan lebih menarik pembeli.
2. Sebuah perusahaan makanan ingin memasarkan produk baru. Namun, mereka belum yakin apakah produknya dapat diterima oleh konsumen atau tidak.

6. Assignment V – Kuis

Kuis Setting Environment

Total points **100/100**

Isilah Pertanyaan di Bawah ini dengan Tepat dan Jujur

✓ Berikut ini yang bukan merupakan code editor/IDE untuk mengoperasikan python adalah *25/25

Jupyter

Excel ✓

Spyder

VSCode

✓ Salah satu kelebihan menggunakan jupyter dibandingkan program lainnya adalah *25/25

Gratis dan mudah diakses

Bisa dipakai tanpa menginstall terlebih dulu

7. Assignment VI – Coordination and Attention to Detail

NAMA : MUHAMMAD ANWAR LIRBOYO

NPM : 201910225319

KELAS : DATA ANALYTICS

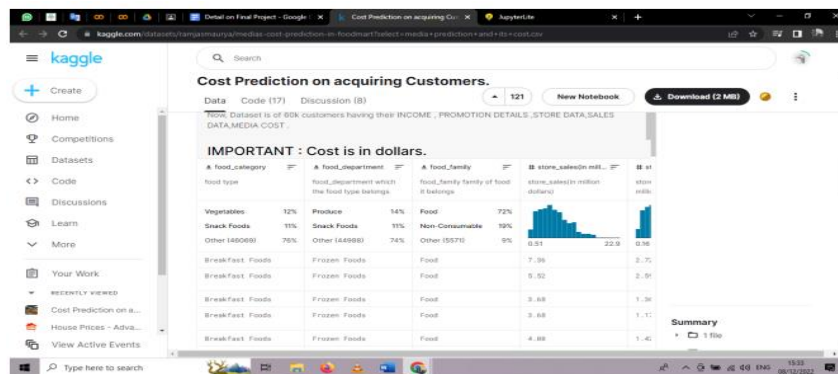
ANALYTICAL THINKING

Isilah pertanyaan dengan jujur dan cermat. Pertanyaan dapat dilihat di dokumen PDF pada tautan di bawah ini. Silahkan tulis jawaban melalui google document atau word, kemudian attach link ataupun file. Good luck!

1. Anda sedang makan siang dengan seorang teman lama dari universitas, dan dia mencari beberapa saran bisnis. Dia sedang berpikir untuk membuka kedai kopi di area Jabodetabek.

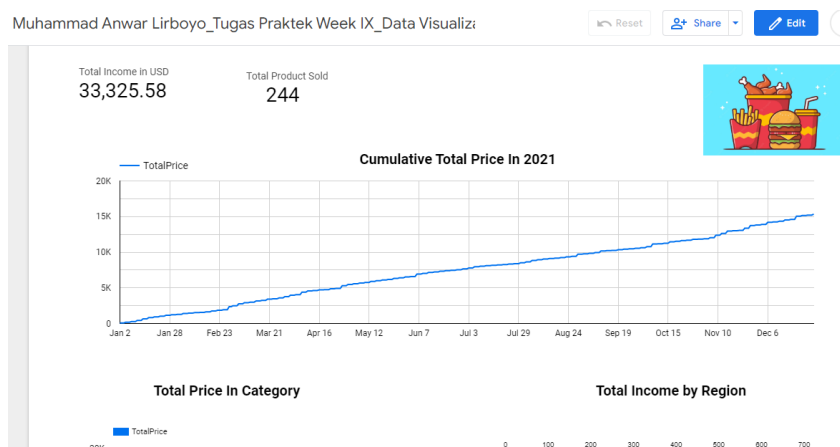
Dia melihat potensi dalam bisnis ini tetapi menginginkan bantuan anda dalam menentukan apakah membuka kedai kopi adalah ide yang bagus.

8. Assignment Business Understanding with Kaggle dataset








9. Assignment Google Data Studio









Muhammad Anwar Lirboyo_Tugas Praktek Week IX_Data Visualiz:



10. Assignment Lainnya di Google Classroom

ndependen Kampus Merdeka Cycle... Stream Classwork People  

 View your work  Google Calendar  Class Drive folder

 FINAL PROJECT	Due Dec 4, 11:59 PM
 Tugas Praktek Week IX - Data Visualization	Due Oct 22, 8:00 AM
 Tugas Praktek Week 8 - Numpy Preprocessi...	Due Oct 14, 11:59 PM
 Tugas Week VII - Statistics with Numpy	Due Oct 10, 11:59 PM
 Tugas Praktek 2 Week VI - Numpy Data Type	Due Oct 4, 11:59 PM
 Tugas Praktek Week VI - Pandas Dataframes	Due Sep 29, 11:59 PM
 Tugas Praktek 2 Week 5 - API	Due Sep 24, 11:59 PM
 Tugas Praktek Week 5 - Working with Text D...	Due Sep 22, 11:59 PM