

Homework

Data Visualization



Estimasi Waktu Pengerjaan



1 - 2 jam

Jumlah Soal



7 Soal

Total Point



100 poin

Teknis Pengerjaan

1. Pekerjaan dilakukan secara **individu**, dengan **menggunakan template yang disediakan**, download [disini](#).
2. Soal-soal berupa **pertanyaan bisnis** yang dijawab dengan **visualisasi**, dan soal-soal dibagi menjadi beberapa bagian berdasarkan tingkat kesulitannya.
 - **Beginner** : 4 soal
 - **Intermediate** : 2 soal
 - **Advance** : 1 soal
3. Upload hasil pengerjaanmu melalui LMS dengan format nama file sebagai berikut **Nama_Lengkap_Batch_XX** dalam format .html (cara save dalam format .html [disini](#))

Dataset yang digunakan

Telco Customer Churn

- **Deskripsi:**

Dataset ini menggambarkan perilaku dan profil pelanggan di perusahaan Telco yang digunakan guna menganalisis dan memprediksi retensi pelanggan.

- **Data:**

Setiap baris mewakili pelanggan, setiap kolom berisi atribut pelanggan.

- **Link download [disini](#)**

Soal *Beginner*

1. Buatlah grafik yang memuat informasi rata-rata biaya bulanan berdasarkan metode pembayarannya! Kemudian tuliskan interpretasi-mu dari grafik tersebut. (10 poin)

Langkah-langkah:

- a. Kelompokkan data berdasarkan metode pembayaran dan gunakan fungsi agregasi yang tepat untuk kasus ini.
- b. Kemudian, buatlah grafik yang tepat menurutmu

Untuk mempermudah kamu, yuk lihat resource di bawah ini:

- Topic Data Processing: Materi DataFrame Aggregation - bagian Groupby



2. Jika masa berlangganan (*tenure*) dikelompokkan menjadi sebagai berikut:

- **low:** User dengan *tenure* < 21 bulan
- **medium:** User dengan *tenure* 21 - 40 bulan
- **high:** User dengan *tenure* > 40 bulan

Buatlah grafik yang menunjukkan jumlah pelanggan berdasarkan pengelompokkan tersebut! (10 poin)

Langkah-langkah

- a. Kelompokkan data dengan menggunakan fungsi `numpy.where` atau `pandas.cut`
- b. Gunakan fungsi agregasi yang tepat untuk kasus ini.
- c. Kemudian, buatlah grafik yang tepat menurutmu

Untuk mempermudah kamu, yuk lihat resource di bawah ini:

- Topic Data Processing: Materi DataFrame Aggregation - bagian Groupby
- Topic Python Programming: Materi Conditional Statement - bagian `np.where()`



10-15 menit

3. Apakah warga senior banyak yang menggunakan layanan telepon? (10 poin)

Langkah-langkah:

- Kelompokkan data berdasarkan dengan apakah pelanggan tersebut merupakan warga senior dan menggunakan layanan telepon, kemudian hitung jumlah unik pelanggannya.
- Buat pula rasio persentase dari jumlah unik customer tersebut.
- Untuk mempermudah, buat grafik dengan data tersebut menggunakan library seaborn dan gunakan parameter 'hue'

Untuk mempermudah kamu, yuk lihat resource di bawah ini:

- Topic Data Processing: Materi DataFrame Aggregation - bagian multi Groupby
- Topic Data Visualization: Materi Advanced Plot - bagian Multi Barplot



4. Bagaimana distribusi dari total biaya? (10 poin)

Langkah-langkah:

- Ubah tipe data kolom total biaya dari *object* menjadi *integer* atau *float* dan perhatikan kelengkapan value dari kolom tersebut.
- Buat plot distribusi untuk kolom tersebut

Untuk mempermudah kamu, yuk lihat resource di bawah ini:

- Topic Data Processing: Materi Transformasi pada Dataframe - bagian Mengubah Tipe Data



Soal *Intermediate*

5. Apakah pelanggan yang memiliki biaya bulanan lebih tinggi akan cenderung untuk berhenti berlangganan (*churn*)? (15 poin)

Langkah-langkah:

- Bandungkan distribusi biaya bulanan antara pelanggan yang berhenti dan tetap berlangganan.
- Untuk mempermudah, buat grafik dengan data tersebut menggunakan library seaborn dan gunakan parameter 'hue'

Untuk mempermudah kamu, yuk lihat resource di bawah ini:

- Topic Data Visualization: Materi Advanced Plot - bagian Multi Distribution plot



6. Bagaimana pengaruh memiliki pasangan & tanggungan terhadap tingkat *churn* pelanggan? (15 poin)

Langkah-langkah:

- a. Gabungkan data yang menunjukkan pelanggan memiliki pasangan dan/atau tanggungan.
 - i. Memiliki pasangan dan tanggungan
 - ii. Memiliki pasangan dan tidak memiliki tanggungan
 - iii. Tidak memiliki pasangan, namun memiliki tanggungan
 - iv. Tidak memiliki pasangan dan tanggungan
- b. Kemudian, hitung rasio pelanggan yang berhenti berlangganan (*churn*)

Untuk mempermudah kamu, yuk lihat resource di bawah ini:

- Topic Data Processing: Materi DataFrame Aggregation - bagian Groupby



Soal *Advanced*

7. Buatlah insight dari data telco customer, dan sertakan storyline pada visualisasi tersebut dengan memanfaatkan atribut tenure, contract, monthly charges, status*, jumlah pelanggan, dan *churn*! (30 poin)

Langkah-langkah:

- Wajib** menggunakan kelima atribut (tenure, contract, monthly charges, status, jumlah pelanggan, *churn*)
- Cari *insight* yang berkaitan dengan kelima atribut tersebut, dan visualisasikan (usahakan membuat visualisasi lebih dari satu)!
- Tambahkan *storyline* di **dalam** visualisasi tersebut

Untuk mempermudah kamu, yuk lihat resource di bawah ini:

- Topic Data Visualization: Materi Advanced Plot - bagian Pengenalan Data Storytelling

*Atribut **status** didapatkan dengan asumsi:

- Married : pelanggan adalah **senior citizen atau memiliki partner atau memiliki dependents**
- Single : tidak memiliki status seperti yang disebutkan diatas

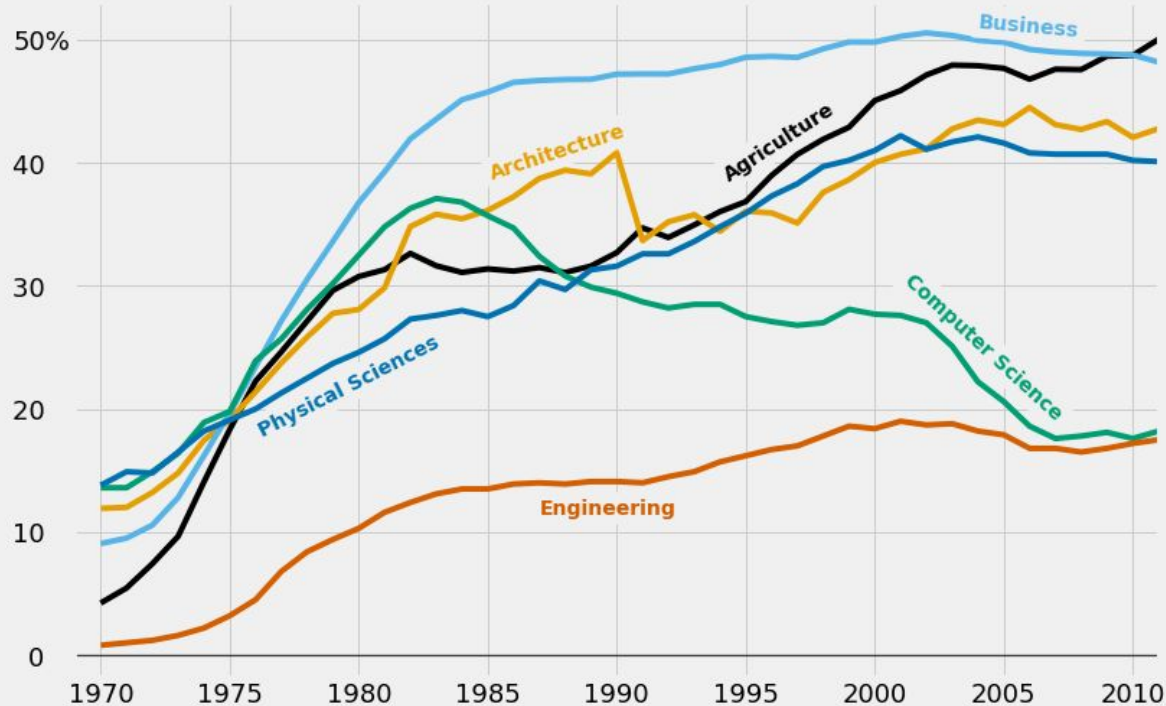


20-35 menit

Contoh!

The gender gap is transitory - even for extreme cases

Percentage of Bachelors conferred to women from 1970 to 2011 in the US for extreme cases where the percentage was less than 20% in 1970



Judul informatif
dan/atau persuasif

Penjelasan atau storyline
terkait visualisasi yang
ditampilkan

- Tutorial untuk membuat plot disamping ([DataQuest](#))
- Contoh Storytelling dengan Python dari [Kaggle](#)

Hint:

Untuk memudahkan bagaimana membuat data storytelling, kamu dapat menjawab beberapa pertanyaan berikut :

Pelanggan **banyak melakukan churn**,

- Pada saat **tenure** berapa bulan?
- Term **contract** jenis apa yang dipilih oleh pelanggan?
- **Status** pelanggan seperti apa? Bagaimana persebarannya jika dilihat dari tenure pelanggan?
- Berapa **monthly charges** pelanggan? Apakah status berpengaruh pada monthly charges pelanggan?

*Pertanyaan diatas hanya sebagai alat bantu saja, kamu dapat **bebas** membuat story telling dengan data yang dimiliki*

Selamat Mengerjakan!