



Homework Guideline



Kita akan menganalisa data dari kaggle berjudul *House Prices* (Harga Rumah)



Search

- Home
- Compete
- Data
- Notebooks
- Discuss
- Courses
- More

Recently Viewed

- House Prices: Advance...
- A/B testing



House Prices: Advanced Regression Techniques

Predict sales prices and practice feature engineering, RFs, and gradient boosting
4,563 teams · Ongoing

Overview **Data** Notebooks Discussion Leaderboard Rules Team My Submissions **Submit Predictions**

Data Description

File descriptions

- train.csv - the training set
- test.csv - the test set
- data_description.txt - full description of each column, originally prepared by Dean De Cock but lightly edited to match the column names used here
- sample_submission.csv - a benchmark submission from a linear regression on year and month of sale, lot square footage, and number of bedrooms

DATA




Download data di sini: (gunakan data train)

<https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data>

Data (199 KB)

API





 `kaggle competitions download -c house-prices-adv...`



 Download All



Data Sources

	sample_submission...	2 columns
	test.csv	80 columns
	train.csv	81 columns
	data_description.txt	

About this file

No description yet

Columns

 Id
SalePrice



Format

1. Save jupyter notebook ke dalam format **.ipynb**



Poin Penilaian

1. Kemampuan Analisa (70%)
 - Sesuai dengan **ekspektasi output**
2. Cara presentasi dalam bentuk grafik/bahasa yang mudah dimengerti tim bisnis - **story telling** (30%)





Part 1. Soal Basic



Soal

Karena ada pernyataan bahwa setiap tahunnya harga rumah itu pasti naik. (tahun terjual : YrSold, harga rumah : SalePrice) Silahkan teman-teman buktikan dengan data yang ada menggunakan ilmu statistika? (Tidak terbatas pada pengujian apapun)

Ekspektasi :

- Penentuan hipotesis
- Pengujian yang digunakan
- Deteksi asumsi beserta cara penanganannya
- Hasil pengujian hipotesis
- Kesimpulan dan interpretasi



Part 2. Soal Advanced



AB Test!



Soal

Kasus yang akan digunakan adalah **membuka akun deposit untuk nasabah bank**. Pada kesempatan kali ini data scientist diminta tolong untuk menganalisa hasil ab-testing suatu perusahaan bank. Apakah **jumlah campaign** mampu meningkatkan keinginan user untuk **membuka akun deposit** atau tidak?

Data yang digunakan berasal dari folder **bank.zip** file yang digunakan adalah **bank-full.csv**

1. Untuk variabel jumlah campaign bisa dicek untuk kolom **campaign**
2. untuk variabel apakah akhirnya user membuka akun atau tidak bisa menggunakan **variabel y**





Ekspektasi Hasil Analisa

1. Pemahaman konsep A/B Testing (Jelaskan sub bagian berikut)
 - a. Apa nama experimentnya?
 - b. Definisikan Hipotesis
 - c. Siapa *participant*-nya?
 - d. Variabel yang akan diuji apa?
 - e. Metrics apa yang akan digunakan?
 - f. Berapa sample size dan durasi experiment?

1. Pengujian Hipotesis
 - b. Penentuan hipotesis
 - c. Pengujian yang digunakan
 - d. Deteksi asumsi beserta cara penanganannya
 - e. Hasil pengujian hipotesis
 - f. Kesimpulan dan interpretasi



**Thank
YOU**

