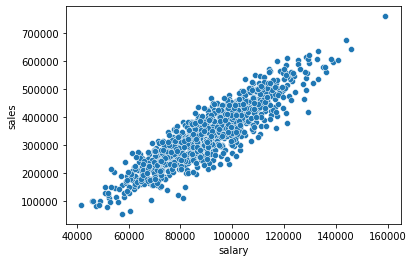
**Test Data Scientist**

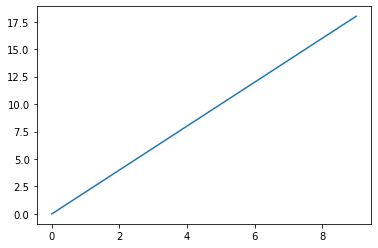
**Test Awal**

1. Sebutkan jenis-jenis grafik yang anda ketahui serta gambarkan bentuknya dengan sederhana!

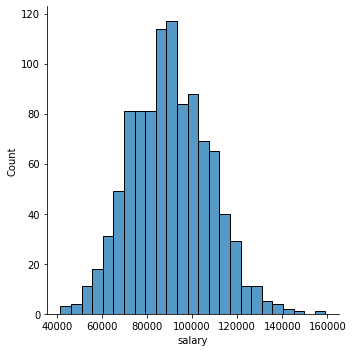
* Scatter plot



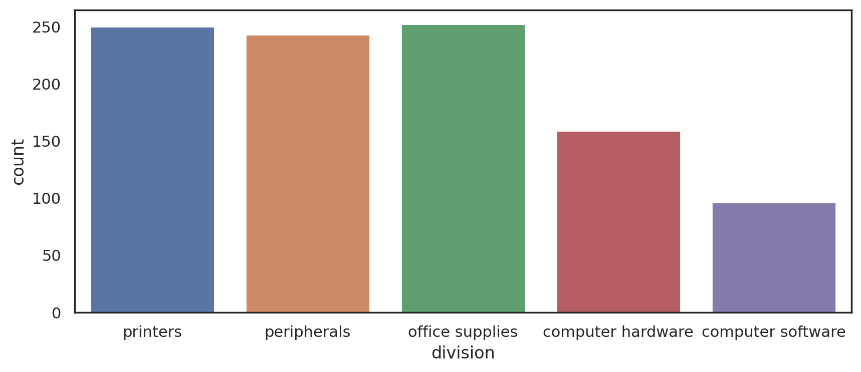
* Line plot



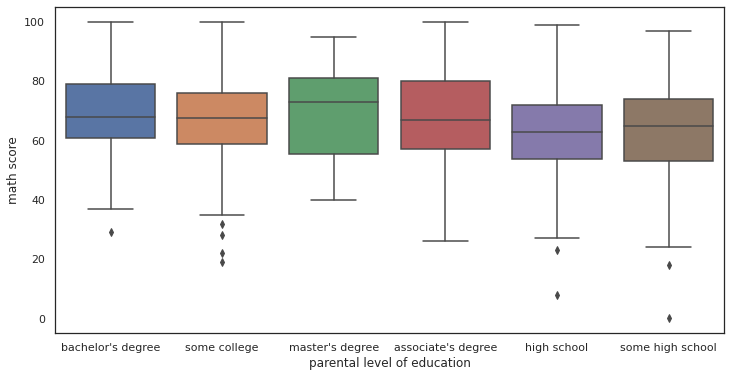
* Histogram plot



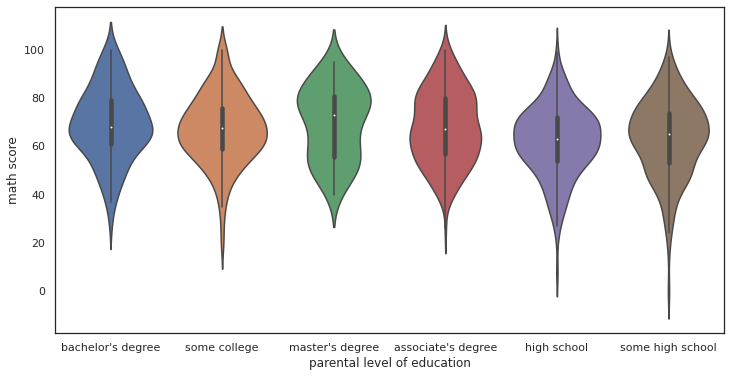
* Bar plot



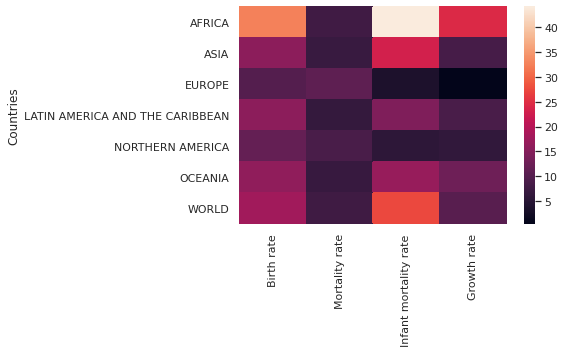
* Box plot



* Violin plot

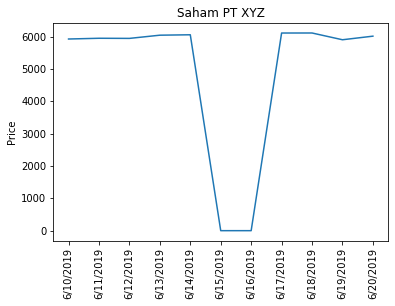


* Heatmap



1. Buatlah visualisasi berdasarkan data yang tersedia dan berikan penjelasan singkat apa yang ingin anda sampaikan lewat visualisasi tersebut!
   1. Data saham PT.XYZ

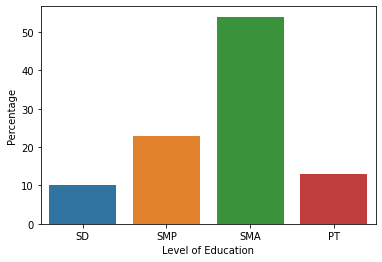
|  |  |
| --- | --- |
| Date | Price |
| 6/10/2019 | 5925 |
| 6/11/2019 | 5948 |
| 6/12/2019 | 5944 |
| 6/13/2019 | 6043 |
| 6/14/2019 | 6056 |
| 6/15/2019 | 0 |
| 6/16/2019 | 0 |
| 6/17/2019 | 6110 |
| 6/18/2019 | 6111 |
| 6/19/2019 | 5902 |
| 6/20/2019 | 6013 |



Berdasarkan grafik saham PT XYZ, dapat diketahui harga saham pada minggu ke-2 di bulan Juni 2019 menyentuh angka 0 pada tanggal 15 dan 16. Tanggal tersebut diketahui adalah weekend sehingga tidak terjadi transaksi saham.

* 1. Data persentase pengguna aplikasi ACK

|  |  |
| --- | --- |
| Level of Education | Percentage (%) |
| SD | 10 |
| SMP | 23 |
| SMA | 54 |
| PT | 13 |



Berdasarkan grafik tersebut, mayoritas pengguna aplikasi ACK memiliki latar belakang pendidikan SMA sedangkan pengguna terendah adalah berpendidikan SD.

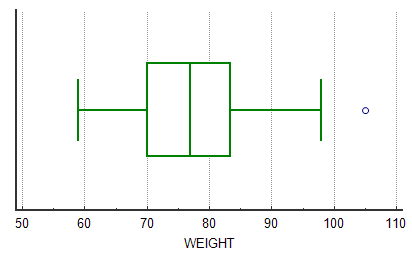
* 1. 5 Negara dengan pengguna Facebox terbesar

|  |  |  |
| --- | --- | --- |
| Country | Active Users (million) | Population (million) |
| Inicia | 294 | 1296 |
| Amenika | 204 | 329 |
| Intoresia | 131 | 262 |
| Bresile | 129 | 208 |
| Mensico | 84 | 126 |

1. a. Berikan interpretasi terhadap plot berikut beserta semua nilai yang digambarkan!

Boxplot tersebut memilki median berkisar pada 77 dan memiliki sebuat outlier

b. Apa yang dapat anda simpulkan dari data tersebut?



1. *Insight* apa saja yang mungkin didapat dari data berikut?

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Id.** | **Tahun Bergabung** | **Jumlah Kiriman** | **Gender** | **Status** | **Domisili** | **Pendidikan** |
| 128474620001 | 2019 | 42 | Laki-laki | Lajang | Jakarta | SMA |
| 128474620002 | 2018 | 43 | Laki-laki | Lajang | Jakarta | Perguruan Tinggi |
| 128474620003 | 2017 | 54 | Laki-laki | Lajang | Bandung | Perguruan Tinggi |
| 128474620004 | 2017 | 12 | Perempuan | Lajang | Jakarta | SMA |
| 128474620005 | 2016 | 14 | Perempuan | Lajang | Surabaya | SMA |
| 128474620006 | 2018 | 112 | Perempuan | Lajang | Jakarta | SMA |
| 128474620007 | 2019 | 1028 | Perempuan | Menikah | Surabaya | SMA |
| 128474620008 | 2019 | 2123 | Laki-laki | Menikah | Bandung | Perguruan Tinggi |
| 128474620009 | 2018 | 2343 | Perempuan | Menikah | Bandung | Perguruan Tinggi |
| 128474620010 | 2018 | 2344 | Laki-laki | Menikah | Bandung | Perguruan Tinggi |
| 128474620011 | 2018 | 2353 | Laki-laki | Menikah | Jakarta | Perguruan Tinggi |
| 128474620012 | 2017 | 1872 | Perempuan | Menikah | Jakarta | SMA |
| 128474620013 | 2018 | 2364 | Perempuan | Menikah | Jakarta | SMA |
| 128474620014 | 2018 | 276 | Perempuan | Menikah | Surabaya | SMA |
| 128474620015 | 2018 | 948 | Perempuan | Lajang | Surabaya | SMA |
| 128474620016 | 2017 | 16 | Laki-laki | Lajang | Bandung | SMA |
| 128474620017 | 2016 | 18 | Perempuan | Lajang | Bandung | SMA |

1. Apa yang anda ketahui tentang *big data*?

Kumpulan yang berukuran sangat besar, kompleks, dan melebihi proses daya tampung dari sistem database yang ada sehingga tidak bisa diproses menggunakan perangkat konvensional.

1. Apa yang anda ketahui mengenai *microtargeting* dan apa hubungannya dengan *big data?*

*Microtargeting* adalah suatu strategi menggunakan data dari masyarakat/pengguna untuk mengetahui siapa mereka, apa yang mereka lakukan, media apa saja yang mereka gunakan dan lain-lain. Data-data tersebut pasti berkaitan dengan *big data,* karena data pengguna tersebut pasti tersimpan dalam suatu database misalnya google, yahoo dan aplikasi lainnya.

1. Sebutkan dan jelaskan tipe-tipe data dalam pemrograman!

* Integer = bilangan bulat
* Float = desimal
* List = kumpulan data sequence

1. Sebutkan dan tuliskan *library* python yang pernah anda gunakan untuk membuat sebuah program!

* Numpy
* Matplotlib
* Seaborn
* Sklearn
* Tensorflow
* Pandas