



Topik Proyek KASDD Ganjil 2023/24

Daftar Topik

1. Property Rent Analysis (PRA)
2. Apartment Listings (AL)
3. Data Delay Penerbangan (DDP)
4. European Investment Management (EIM)
5. Valorant Pro Matches (VPM)
6. Spesifikasi Motor (SM)
7. Telco Customer Behavior (Marketing)
8. Ultimate Fighter Championship (UFC)
9. Customer Credit Loan (CCL)
10. Meramal Cuaca (MC)

Dataset dari seluruh topik akan segera menyusul

Pengajuan Topik

- Pengajuan topik dilakukan pada kuesioner yang ada di laman SCELE.
- Setiap kelompok wajib mengajukan **3 topik yang diprioritaskan** pada kuesioner yang ada di scele, kemudian tim asisten akan memilih 1 topik dari ketiga topik tersebut.
- Pengajuan topik bersifat FCFS (*First Come First Serve*). Ketika suatu topik sudah dipilih oleh kelompok lain, maka secara otomatis topik tersebut tidak dapat dipilih lagi. Oleh sebab itu teman-teman diminta memilih beberapa topik berbeda yang diprioritaskan. Jika ketiga topik sudah terisi oleh kelompok lain, maka topik akan di acak oleh tim asisten.
- Bagi perwakilan kelompok yang tidak mengajukan topik hingga batas waktu yang ditentukan, maka tim asisten akan memilih topik secara acak.
- Jika ada kendala pengajuan topik, harap segera hubungi asisten dosen.
- **Tidak ada toleransi atas keterlambatan dan keputusan final tidak dapat diganggu gugat**

01

Property Rent Analysis



Deskripsi Topik 01

Deskripsi Dataset	Dataset ini berisi informasi tentang iklan sewa properti di sekitar wilayah Dubai dan Abu Dhabi, Uni Emirat Arab.
Jumlah Kolom x Baris	13 kolom x 18450 baris
Deskripsi Kolom	

- id: ID unik dari iklan
- emirates: Emirates dari properti
- established_year: Tahun pendirian properti
- furnish_status: Jenis furnish dari properti
- monthly_rent_aed: Biaya sewa per bulan (dalam Uni Emirat Arab Dirham)
- num_of_bathroom: Jumlah kamar mandi pada properti
- num_of_parking_space: Jumlah lahan parkir pada properti
- num_of_rooms: Jumlah ruangan pada properti
- property_facilities: Fasilitas dari properti
- property_location: Lokasi dari properti (emirates - neighborhoods)
- property_name: Nama dari properti
- property_size: Luas area dari properti (dalam square feet)
- property_type: Tipe dari properti

Deskripsi Topik 01

Gambaran Dataset

id	emirates	established_year	furnish_status	monthly_rent_aed	num_of_bathroom	num_of_parking_spa	num_of_rooms	property_facilities	property_location	property_name	property_size	property_type
0	Dubai		Fully Furnished	1638	2	1	3	Parking, Security, Sv	Dubai - Jumeirah	Danau Kota Suite	829	Condominium
1	Dubai	2021	Partially Furnished	1794	2	1	2	Playground, Security	Dubai - Al Wasl	The Era	1020	Condominium
2	Dubai		Fully Furnished	1287	2	1	3	Parking, Air-Cond, C	Dubai - Jumeirah	Kenari Apartment (T	900	Apartment
3	Abu Dhabi	2022	Not Furnished	935.22	2	2	2	Parking, Security, Lif	Abu Dhabi - Al Hebia	Vertex Tower @ Cyb	700	Service Residence
4	Dubai		Not Furnished	1248	2	1	3	Parking, Security, Lif	Dubai - Rigga Al But	Tuan Residency	878	Service Residence
5	Dubai		Partially Furnished	1014	2		3	Parking, Security, Ai	Dubai - Al Rigga	Perdana Villa (Pand	900	Apartment
6	Abu Dhabi	2016	Fully Furnished	936	1	1	1	Parking, Security, Lif	Abu Dhabi - Al Hebia	Solstice @ Pan'gaea	450	Condominium
7	Dubai	2007	Fully Furnished	936	1		1		Dubai - Al Khawanee	Maytower	351	Service Residence
8	Abu Dhabi	2006	Partially Furnished	975	2	1	4	Jogging Track, Barbe	Abu Dhabi - Al Mery	Greenview Residenc	1366	Condominium
9	Abu Dhabi		Not Furnished	936	2	2	3	Parking, Security, Lif	Abu Dhabi - Al Mery	Vista Sungai Ramal	1000	Apartment
10	Abu Dhabi	2022	Fully Furnished	546	1		1	Playground, Swimm	Abu Dhabi - Nadd Al	Horizon Suites	240	Others
11	Abu Dhabi		Fully Furnished	834.6	2	2	3	Parking, Security, Jo	Abu Dhabi - Al Mank	PANGSAPURI SRI P	865	Apartment
12	Abu Dhabi	2017	Not Furnished	897	2		3	Playground, Gymnas	Abu Dhabi - Al Mam	Kiara Plaza	1097	Service Residence
13	Abu Dhabi	2015	Partially Furnished	858	1	2	1	Security, Lift, Minim	Abu Dhabi - Me'aise	Empire City	952	Condominium
14	Dubai	2021	Partially Furnished	1248	2	1	2	Playground, Parking	Dubai - Al Rigga	Majestic Maxim	670	Service Residence
15	Abu Dhabi	2018	Partially Furnished	1014	2	2	3	Security, Jogging Tr	Abu Dhabi - Al Mery	Oasis 1 @ Mutiara H	1250	Condominium
16	Abu Dhabi	2021	Fully Furnished	858	1	1	1	Security, Playgrounc	Abu Dhabi - Graytee	DK Impian	592	Service Residence
17	Abu Dhabi		Partially Furnished	1248	2	2	3	Lift, Multipurpose ha	Abu Dhabi - Al Mery	Residensi Sutera 7	1120	Condominium
18	Dubai	1999	Partially Furnished	1092	2		3	Playground, Gymnas	Dubai - Al Rigga	Bayu Task 1	920	Condominium
19	Dubai		Fully Furnished	1560	2	1	3	Minimart, Gymnas	Dubai - Al Murar	Vista Komanwel	1422	Condominium
20	Dubai	2009	Fully Furnished	1365	2	1	2	Swimming Pool, Gyr	Dubai - Port Saeed	Viva Residency	900	Condominium
21	Dubai	2009	Fully Furnished	1559.22	2	1	3	Swimming Pool, Pla	Dubai - Jumeirah Isl	Changkat View Cond	1000	Condominium
22	Dubai	2017	Fully Furnished	1287	2	1	3	Security, Parking, Mi	Dubai - Zabeel	Seasons Garden Res	770	Apartment
23	Dubai		Partially Furnished	1950	3		4	Parking, Security, Lif	Dubai - Al Baraha	D'Suria Condo Tama	1260	Condominium
24	Dubai		Fully Furnished	819	2	1	3	Parking, Security, Pl	Dubai - Oud Al Mute	Dahlia Apt	806	Apartment
25	Dubai		Not Furnished	936	1		3	Parking, Playground	Dubai - Jumeirah	Flat Taman Melati	900	Others
26	Abu Dhabi	2016	Fully Furnished	936	1		1	Parking, Security, Lif	Abu Dhabi - Al Hebia	Solstice @ Pan'gaea	450	Condominium
27	Dubai		Not Furnished	429	1		3		Dubai - Al Qusais	taman intan baiduri	600	Others
28	Dubai		Partially Furnished	1326	2	1	3	Swimming Pool, Par	Dubai - Zabeel	Fera Residence @ T	688	Service Residence
29	Dubai	2014	Fully Furnished	2028	3	2	4	Security, Barbeque	Dubai - Al Twar	Arte KL	1500	Condominium

02

Apartment Listings



Problem Guideline Topik 01

1. Lakukan eksplorasi sederhana pada data dan ceritakan insight yang Anda dapatkan dari data tersebut! (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Apakah terdapat hubungan khusus antara status furnish suatu properti dengan fitur lain pada properti tersebut?
 - b. Bagaimana ciri-ciri properti yang memiliki tipe condominium dan fasilitas jogging track?
 - c. Apa fasilitas properti yang paling mempengaruhi harga sewa?
 - d. Bagaimana persebaran neighborhoods berdasarkan rata-rata harga sewa dan luas properti?

Mahasiswa perlu menambahkan **dua poin eksplorasi sendiri**

2. Buatlah beberapa model yang mungkin akan diperlukan, di antaranya: (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Buatlah model untuk mengklasifikasi tipe dari suatu properti!
 - b. Buatlah model untuk memprediksi luas area dari suatu properti!
 - c. Lakukan clustering pada dataset dan hubungkan karakteristik dari masing-masing cluster yang didapatkan dengan suatu fitur, seperti fasilitas properti atau tipe properti!

Deskripsi Topik 02

Deskripsi Dataset	Dataset ini berisikan berbagai informasi suatu apartemen. Dataset ini mencakup kondisi sekitar apartemen dan apartemen sendiri serta tawaran pada apartemen.
Jumlah Kolom x Baris	19 kolom x 12.2k baris
Deskripsi Kolom	

- ID: ID yang unik untuk setiap apartemen.
- latitude: Nilai latitude (garis lintang) untuk apartemen yang sesuai.
- longitude: Nilai longitude (garis bujur) untuk apartemen yang sesuai.
- district: Distrik atau wilayah di mana apartemen berada.
- address_offers: Jumlah tawaran atau penawaran yang terkait dengan alamat apartemen.
- agent_offers: Jumlah tawaran atau penawaran yang terkait dengan agen properti yang berhubungan dengan apartemen.
- subway_offers: Jumlah tawaran atau penawaran yang terkait dengan stasiun kereta yang berdekatan dengan apartemen.
- closest_subway: Stasiun kereta terdekat dengan apartemen.
- dist_to_subway: Jarak antara apartemen dan stasiun kereta terdekat.
- subway_grade: Kelas stasiun kereta yang berdekatan dengan apartemen.
- subway_dist_to_center: Jarak antara stasiun kereta terdekat dan pusat kota.
- rooms: Jumlah kamar di apartemen.
- floor: Lantai tempat apartment berada.
- max_floor: Jumlah lantai maksimum dalam bangunan di mana apartemen berada.
- footage: Luas dari apartemen.
- material: Informasi tentang bahan bangunan apartemen.
- repair: Kondisi perbaikan apartemen.
- category_age: Kategori usia apartemen.
- price: Harga apartemen.

Deskripsi Topik 02

Gambaran Dataset

ID	latitude	longitude	district	address_offers	agent_offers	subway_offers	closest_subway	dist_to_subway	subway_grade	subway_dist_to_cent	rooms	floor	max_floor	footage	material	repair	category_age	price
MAL0	55.609009	37.734632	UAO		1		Park Pobedy		Underground	17166	One Room	11		39.4	panel	Good	Old buildings	8000000
MAL1	55.60876	37.741936		1	1	9	Park Pobedy	802	Underground	17166	One Room		9	35.6	panel	Enough	Old buildings	7200000
MAL2	55.6102066		UAO		2	57	Park Pobedy	638	Underground	17166	Two Room	5		53	panel	Enough	Old buildings	10300000
MAL3	55.6240169	37.7403735	UAO		1	292	Park Pobedy	929	Underground	17166	Two Room	8	9	44.1	panel		Old buildings	8500000
MAL4	55.6091073	37.74043	UAO		1		5 Park Pobedy	785	Underground		Two Room	6	9		panel	Enough	Old buildings	9000000
MAL5	55.6222942	37.7335547	UAO		1	292	5 Park Pobedy	962	Underground	17166	One Room	1	12	39.5	panel	Enough	Middle-aged building	8050000
MAL6	55.6208204	37.7422834	UAO		1		9 Park Pobedy	556	Underground		Two Room	19	22	57	panel	Enough	Middle-aged building	10900000
MAL7	55.62362	37.74428	UAO		1	36	13 Park Pobedy		Underground	17166	One Room	2		33.9	block		Middle-aged building	8100000
MAL8	55.622705	37.726583			1		6 Park Pobedy		Underground	17166	Two Room	10	12	45.9	panel	Good	Old buildings	10000000
MAL9		37.732072	UAO		1			796	Underground	17166	Two Room	6	12		panel	Good	Old buildings	9000000
MAL10	55.6102066	37.7424953	UAO			2	8 Park Pobedy	638	Underground	17166	Two Room	8	12		panel	Poor	Old buildings	
MAL11		37.73667			1	218	9	529	Underground	17166	Two Room			51	panel	Enough	Old buildings	10000000
MAL12		37.7412443	UAO		1		Park Pobedy	433	Underground	17166	Two Room	13		53.5	panel	Enough	Old buildings	10800000
MAL13	55.6222942	37.7335547	UAO		1	205	5 Park Pobedy	962		17166	One Room	1	12	39.5	panel		Middle-aged building	7850000
MAL14	55.6191015	37.7574297			1	205	5 Park Pobedy	923	Underground	17166	One Room	2	14	39	panel	Enough	Middle-aged building	7800000
MAL15	55.62386	37.73837			1	205	Park Pobedy	950	Underground	17166	One Room	4	9	32.9	panel	Enough	Old buildings	8000000
MAL16	55.612197	37.737713	UAO		1	35	8 Park Pobedy	564	Underground	17166	One Room		12	38.1	panel	Good	Middle-aged building	8500000
MAL17	55.6145286	37.7396051			1	1		307	Underground	17166	Two Room	5	9	44	panel		Middle-aged building	9000000
MAL18	55.6232024	37.7482036	UAO		1	2	Park Pobedy	857	Underground	17166	Two Room		12		panel		Old buildings	10000000
MAL19	55.6222942	37.7335547	UAO		2	5	9 Park Pobedy	962	Underground	17166	One Room	1	12	39.5	panel	Good	Middle-aged building	7850000
MAL20	55.6091073	37.74043	UAO		1	11	6 Park Pobedy	785	Underground	17166	Two Room		9		panel	Enough	Old buildings	8999999
MAL21	55.621724	37.756632	UAO		1	1	9 Park Pobedy	1031	Underground	17166	One Room	2	14	39	panel	Enough	Middle-aged building	7700000
MAL22	55.6102066	37.7424953	UAO		1	1	6 Park Pobedy	638	Underground	17166	Two Room	5	12	53	panel	Enough	Old buildings	10300000
MAL23		37.7395941	UAO		1		8 Park Pobedy		Underground		Two Room		12		panel	Good	Old buildings	9000000
MAL24		37.752436	UAO		1	62	8 Park Pobedy	539	Underground	17166	Two Room	3	9		panel	Good	Old buildings	9800000
MAL25	55.62362	37.74428	UAO		1	7	9 Park Pobedy	859	Underground	17166		2	22	33.9	block	Enough	Middle-aged building	8000000
MAL26		37.752436			1	2	13 Park Pobedy	539	Underground	17166	Two Room	8	9		panel	Enough	Old buildings	9000000
MAL27		37.751314	UAO		1	1	Park Pobedy	483	Underground	17166	Two Room		9		panel	Enough	Old buildings	8600000
MAL28	55.615182	37.743148	UAO		1	12	8 Park Pobedy		Underground	17166	One Room	3	22	38.6	panel		Middle-aged building	8600000
MAL29	55.6091073	37.74043	UAO		2		8 Park Pobedy	785	Underground	17166	Two Room	6	9	45	panel	Good	Old buildings	8990000
MAL30	55.609406	37.741136	UAO		1	4	Park Pobedy	741	Underground		Two Room	9	9		panel	Enough	Old buildings	8500000
MAL31	55.615182	37.743148	UAO		1		11 Park Pobedy	91	Underground	17166		3		38.4	panel	Good	Middle-aged building	8500000
MAL32	55.62386	37.73837	UAO		1		13 Park Pobedy	950	Underground	17166	One Room		9		panel		Old buildings	8000000
MAL33	55.61643	37.739323	UAO		1	9	11	291		17166	One Room		9	32.5	panel	Enough	Old buildings	7190000

Problem Guideline Topik 02

1. Lakukan eksplorasi sederhana pada data dan ceritakan insight yang Anda dapatkan dari data tersebut! (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Apakah terdapat hubungan antara lokasi bangunan dengan harga bangunan?
 - b. Apa ciri-ciri bangunan yang dekat dengan subway dengan jenis kereta yang bukan bawah tanah?
 - c. Apakah kondisi perbaikan bangunan yang buruk merupakan bangunan tua?
 - d. Apakah distrik dengan total tawaran tertinggi?

Mahasiswa perlu menambahkan **dua poin eksplorasi sendiri**

2. Buatlah beberapa model yang mungkin akan diperlukan, di antaranya: (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Buatlah model untuk mengklasifikasi kategori usia bangunan dari apartemen
 - b. Buatlah model untuk memprediksi harga jual dari apartemen
 - c. Lakukan clustering pada dataset, dan berikan analisis dari jenis-jenis apartemen yang terdapat dalam suatu cluster!

03

Data Delay Penerbangan



Deskripsi Topik 03

Deskripsi Dataset	Dataset ini merupakan data penerbangan yang delay pada tahun 2022 di bulan januari. Dataset ini mengandung atribut status dan waktu penerbangan yang terjadi.
Jumlah Kolom x Baris	35 kolom x 55.449 baris
Deskripsi Kolom	

- Year: Tahun dari catatan data penerbangan.
- Quarter: Kuartal tahun ketika penerbangan berlangsung
- Month: Bulan ketika penerbangan berlangsung
- DayofMonth: Hari dalam bulan ketika penerbangan berlangsung (1-31).
- DayOfWeek: Hari dalam seminggu ketika penerbangan berlangsung (1=Senin, 2=Selasa, ..., 7=Minggu).
- FlightDate: Tanggal ketika penerbangan berlangsung dalam format YYYY-MM-DD.
- Marketing_Airline_Network: Jaringan atau grup maskapai pemasaran.
- Operated_or_Branded_Code_Share_Partners: Informasi tentang mitra atau maskapai yang terlibat dalam operasi.
- DOT_ID_Marketing_Airline: Identifier Departemen Transportasi (DOT) dari maskapai pemasaran.
- IATA_Code_Marketing_Airline: Kode IATA dari maskapai pemasaran.
- Flight_Number_Marketing_Airline: Nomor penerbangan dari maskapai pemasaran.
- Operating_Airline: Maskapai yang mengoperasikan penerbangan.
- DOT_ID_Operating_Airline: Identifier Departemen Transportasi (DOT) dari maskapai yang mengoperasikan penerbangan.
- IATA_Code_Operating_Airline: Kode IATA dari maskapai yang mengoperasikan penerbangan.
- Tail_Number: Nomor ekor pesawat.
- Flight_Number_Operating_Airline: Nomor penerbangan dari maskapai yang mengoperasikan penerbangan.
- OriginAirportSeqID: Identifier urutan bandara asal.
- Origin: Bandara asal penerbangan.

Deskripsi Topik 03

Deskripsi Kolom

- Dest: Tujuan penerbangan.
- CRSDepTime: Waktu keberangkatan yang direncanakan (dalam format HHMM).
- DepTime: Waktu sebenarnya pesawat berangkat (dalam format HHMM).
- TaxiOut: Waktu yang dibutuhkan pesawat di darat untuk taxi keluar dari gerbang.
- WheelsOff: Waktu pesawat memasuki bannya (dalam format HHMM).
- WheelsOn: Waktu pesawat mengeluarkan bannya (dalam format HHMM).
- TaxiIn: Waktu yang dibutuhkan pesawat di darat untuk taxi masuk ke gerbang.
- CRSArrTime: Waktu kedatangan yang direncanakan (dalam format HHMM).
- ArrTime: Waktu sebenarnya pesawat tiba (dalam format HHMM).
- Cancelled: Status penerbangan yang dibatalkan (0=Belum dibatalkan, 1=Dibatalkan).
- Diverted: Status penerbangan yang dialihkan (0=Tidak dialihkan, 1=Dialihkan).
- CRSElapsedTime: Waktu penerbangan yang direncanakan (dalam menit).
- ActualElapsedTime: Waktu penerbangan sebenarnya (dalam menit).
- AirTime: Waktu yang dihabiskan pesawat di udara (dalam menit).
- Flights: Jumlah penerbangan.
- Distance: Jarak penerbangan (dalam mil).
- DelayReason: Alasan utama keterlambatan penerbangan.

Deskripsi Topik 03

Gambaran Dataset

Year	Quarter	Month	DayofMor	DayOfWei	FlightDate	Marketing	Operated	DOT_ID_N	IAATA_Cod	Flight_Nu	Operating	DOT_ID_C	IAATA_Cod	Tail_Num	Flight_Nu	OriginAirr	Origin	Dest	CRSDepTi	DepTime	TaxiOut	WheelsOf	WheelsOr	TaxiIn	CRSArrTin	ArrTime	Cancelled	Diverted	
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	1672	DL	19790	DL	N6704Z	1672	1039707	ATL	JAX	1633	1736	12	1748	1830	4	1738	1834	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	1696	DL	19790	DL	N124DX	1696	1169706	FLL	ATL	1820	1930	13	1943	2106	8	2016	2114	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	1809	DL	19790	DL	N386DA	1809	1039707	ATL	STX	1000	1126	13	1139	1554	3	1440	1557	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	1809	DL	19790	DL	N386DA	1809	1502704	STX	ATL	1555	1656	11	1707	1938	11	1913	1949	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2056	DL	19790	DL	N846DN	2056	1452401	RIC	ATL	1755	1906	10	1916	2040	5	1944	2045	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2183	DL	19790	DL	N397DN	2183	1348702	MSP	TPA	1250	1501	9	1510	1850	5	1659	1855	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2185	DL	19790	DL	N593NW	2185	1463502	RSW	DTW	635	715	14	729	943	7	929	950	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2189	DL	19790	DL	N971AT	2189	1039707	ATL	HSV	1754	1838	10	1848	1820	4	1750	1824	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2223	DL	19790	DL	N355NW	2223	1039707	ATL	ORF	1225	1339	10	1349	1450	5	1355	1455	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2254	DL	19790	DL	N986AT	2254	1323202	MDW	ATL	600	705	14	719	945	10	858	955	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2270	DL	19790	DL	N339NB	2270	1348702	MSP	DFW	2115	2203	33	2236	26	7	2358	33	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2273	DL	19790	DL	N329NB	2273	1348702	MSP	RDU	1940	2101	27	2128	22	7	2322	29	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2287	DL	19790	DL	N665DN	2287	1039707	ATL	TPA	1958	2040	14	2054	2154	4	2125	2158	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2295	DL	19790	DL	N358NW	2295	1039707	ATL	OMA	1959	2027	10	2037	2144	6	2128	2150	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2308	DL	19790	DL	N330NB	2308	1449202	RDU	MCO	1940	2135	16	2151	2312	8	2131	2320	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2321	DL	19790	DL	N6704Z	2321	1039707	ATL	TPA	2259	2337	12	2349	48	4	25	52	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2323	DL	19790	DL	N110DU	2323	1474703	SEA	PHX	1939	2045	18	2103	23	3	2331	26	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2330	DL	19790	DL	N774DE	2330	1410702	PHX	SEA	1210	1905	10	1915	2056	6	1427	2102	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2338	DL	19790	DL	N349DX	2338	1039707	ATL	DFW	2100	2144	11	2155	2243	8	2224	2251	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2361	DL	19790	DL	N376NW	2361	1039707	ATL	OMA	1616	1727	10	1737	1841	5	1742	1846	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2372	DL	19790	DL	N103DU	2372	1486903	SLC	SAT	1000	1102	12	1114	1426	4	1343	1430	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2379	DL	19790	DL	N839MH	2379	1039707	ATL	SAN	1655	1927	15	1942	2045	4	1844	2049	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2393	DL	19790	DL	N539US	2393	1143302	DTW	LAX	1255	1455	12	1507	1616	7	1500	1623	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2457	DL	19790	DL	N389DA	2457	1474703	SEA	ANC	2135	2210	15	2225	35	6	23	41	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2464	DL	19790	DL	N354NB	2464	1348702	MSP	SAT	1026	1120	10	1130	1349	5	1329	1354	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2470	DL	19790	DL	N130DU	2470	1486903	SLC	SFO	1715	1757	14	1811	1850	5	1823	1855	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2488	DL	19790	DL	N538US	2488	1039707	ATL	MCO	2339	31	12	43	139	4	104	143	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2505	DL	19790	DL	N133DU	2505	1129806	DFW	SLC	1835	1918	19	1937	2056	5	2037	2101	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2537	DL	19790	DL	N370NW	2537	1142307	DSM	ATL	1835	2050	14	2104	2328	9	2154	2337	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2588	DL	19790	DL	N112DN	2588	1148102	ECP	ATL	731	825	9	834	1017	6	940	1023	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2625	DL	19790	DL	N353NB	2625	1348702	MSP	MKE	1631	1752	14	1806	1856	5	1748	1901	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2676	DL	19790	DL	N671DN	2676	1039707	ATL	TPA	2059	2149	17	2206	2306	5	2223	2311	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2682	DL	19790	DL	N3737C	2682	1039707	ATL	CVG	2234	2306	13	2319	18	5	2359	23	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2693	DL	19790	DL	N900PC	2693	1039707	ATL	JAX	2245	2326	20	2346	30	4	2354	34	0	0
2022	1	1	1	6	4	1/6/2022	DL	DL	19790	DL	2695	DL	19790	DL	N321US	2695	1226603	IAH	ATL	1557	1746	16	1802	2037	6	1856	2043	0	0

Problem Guideline Topik 03 (1)

1. Lakukan eksplorasi sederhana pada data dan ceritakan insight yang Anda dapatkan dari data tersebut! (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Berapa rata-rata lama delay pada pesawat di setiap faktor penyebab utama dari delay?
 - b. Apa alasan paling umum penyebab pesawat itu delay dan kapan waktu seringnya terjadi delay (pagi/siang/sore/malam)? dan apakah sama untuk setiap maskapai dan aiport?
 - c. Identifikasikan faktor-faktor yang menyebabkan delay pada penerbangan domestik dan internasional di bulan januari 2022. Selain itu, jelaskan apakah ada perbedaan dalam dampak faktor-faktor tersebut antara penerbangan domestik dan internasional, dan apakah ada interaksi antara faktor-faktor tersebut yang memengaruhi tingkat delay.
 - d. Berikan 10 peringkat teratas untuk maskapai apa dan rute mana yang paling sering terbang selama bulan januari 2022

Mahasiswa perlu menambahkan **dua poin eksplorasi sendiri**

Problem Guideline Topik 03 (2)

2. Buatlah beberapa model yang mungkin akan diperlukan, di antaranya: (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
- a. Buatlah model untuk mengklasifikasi alasan delay dari penerbangan (delayReason)
 - b. Buatlah model untuk memprediksi berapa lama pesawat berjalan di darat di setiap penerbangan
 - c. Lakukan clustering pada dataset, dan berikan analisis dari jenis-jenis penerbangan yang terdapat dalam suatu cluster!

04

European Investment Management



Deskripsi Topik 04

Deskripsi Dataset	Dataset ini terdiri dari data performa ETF dan reksa dana yang beroperasi di Eropa.
Jumlah Kolom x Baris	117 kolom x 28.025 baris
Deskripsi Kolom	

- Ticker: Kode unik
- Category: Kategori
- Rating: Penilaian (1-5) yang diberikan Morningstar
- Dividend_frequency: Periode pembagian dividen
- Equity_style: Gaya pengelolaan investasi
- Equity_size: Golongan nilai ekuitas
- Equity_size_score: Skor nilai ekuitas
- Price_prospective_earnings:
- Price_book_ratio: Rasio harga dengan nilai buku
- Price_sales_ratio: Rasio harga dengan pendapatan
- Price_cash_flow_ratio: Rasio harga dengan arus kas
- Dividend_yield_factor: % dividen dibandingkan dengan harga
- Long_term_projected_earnings_growth: % proyeksi pertumbuhan laba jangka panjang
- Historical_earnings_growth: % pertumbuhan laba secara historis
- Sales_growth: % pertumbuhan pendapatan
- Book_value_growth: % pertumbuhan nilai buku
- Roa: % laba dibandingkan dengan nilai asset
- Roe: % laba dibandingkan dengan nilai ekuitas
- Roic: % laba dibandingkan dengan nilai yang diinvestasikan

Deskripsi Topik 04

Deskripsi Kolom

- Bond_interest_rate_sensitivity:
- Bond_credit_quality:
- Average_coupon_rate:
- Average_credit_quality:
- Modified_duration:
- Effective_maturity:
- Asset_stock: % saham dari total portofolio
- Asset_bond: % obligasi dari total portofolio
- Asset_cash: % kas dari total portofolio
- Asset_other: % selain saham, obligasi, dan kas dari total portofolio
- Sector_basic_materials:
- Sector_consumer_cyclical:
- Sector_financial_services
- Sector_real_estate
- Sector_consumer_defensive:
- Sector_healthcare:
- Sector_utilities:
- Sector_communication_services:
- Sector_energy:
- Sector_industrials:
- Sector_technology:
- Market_cap_giant:
- Market_cap_large:
- Market_cap_medium:
- Market_cap_small:
- Market_cap_micro:

Deskripsi Topik 04

Deskripsi Kolom

- Credit_aaa
- Credit_aa
- Credit_a
- Credit_bbb
- Credit_bb
- Credit_b
- Credit_below_b
- Credit_not_rated
- Holdings_n_stocks
- Holdings_n_bonds
- Ongoing_cost
- Management_fees
- Enviromental_score
- Social_score
- Governance_score
- Sustainability_score
- Sustainability_rank
- Sustainability_percentage_rank
- Involvement_abortive_contraceptive
- Involvement_alcohol
- Involvement_animal_testing
- Involvement_controversial_weapons
- Involvement_gambling
- Involvement_gmo
- Involvement_military_contracting
- Involvement_nuclear

Deskripsi Topik 04

Deskripsi Kolom

- involvement_palm_oil
- involvement_pesticides
- involvement_small_arms
- involvement_thermal_coal
- involvement_tobacco
- nav_per_share_currency
- nav_per_share
- shareclass_size_currency
- shareclass_size
- fund_size_currency
- fund_size
- fund_trailing_return_ytd
- fund_trailing_return_3years
- fund_trailing_return_5years
- fund_trailing_return_10years
- fund_return_2019
- fund_return_2018
- fund_return_2017
- fund_return_2016
- fund_return_2015
- fund_return_2020_q3
- fund_return_2020_q2
- fund_return_2020_q1
- fund_return_2019_q4
- fund_return_2019_q3
- fund_return_2019_q2

Deskripsi Topik 04

Deskripsi Kolom

- fund_return_2019_q1
- fund_return_2018_q4
- fund_return_2018_q3
- fund_return_2018_q2
- fund_return_2018_q1
- fund_return_2017_q4
- fund_return_2017_q3
- fund_return_2017_q2
- fund_return_2017_q1
- fund_return_2016_q4
- fund_return_2016_q3
- fund_return_2016_q2
- fund_return_2016_q1
- fund_return_2015_q4
- fund_return_2015_q3
- fund_return_2015_q2
- fund_return_2015_q1
- quarters_up
- quarters_down

Deskripsi Topik 04

Gambaran Dataset

ticker	category	rating	dividend_frequency	equity_style	equity_size	equity_size_score	price_prospective_e	price_book_ratio	price_sales_ratio	price_cash_flow_rat	dividend_yield_fact	long_term_projectec	historical_earnings_e	sales_growth	cash_flow_growth	book_value_growth	roa
0P00000C5S	Switzerland Equity		3 Annually	Blend	Large	328.15	15.73	2.46	2.21	14.8	3.23	6.99	6.24	1.25	5.94	-1.79	8.23
0P00000FY	Sweden Equity		2	Blend	Large	248.09	14.9	1.87	1.77	10.26	1.45	7.6	8.97	2.66	7.23	4.43	7.4
0P00000GCE	Sector Equity Consu		3 Quarterly	Blend	Large	308.52	17.64	2.54	2.38	14.41	2.59	4.95	-10.14	5.86	6.15	18.05	5.7
0P00000HSN	Eurozone Large-Cap		4 Annually	Blend	Large	275.97	14.34	1.32	0.97	7.89	2.6	10.44	-1.07	0.7	5.49	4.09	3.53
0P00000H5W	Japan Large-Cap Eq		3 Annually	Blend	Large	286.08	17.8	1.19	0.88	8.4	2.75	9.43	-3.99	3.52	2.95	4.93	4.72
0P00000HUH	Global Large-Cap Bl		5 Annually	Growth	Large	381.38	21.84	4.35	3.5	14.7	1.88	9.64	10.91	7.5	8.32	4.54	11.07
0P00000HUX	US Large-Cap Blend		3 Annually	Value	Large	314.59	18.15	3.22	2.09	13.61	2.46	7.56	13.87	6.41	-1.91	5.91	8.65
0P00000IQX	UK Large-Cap Equity		3 Quarterly	Value	Large	266.54	12.59	1.31	1.01	7.49	4.08	6.63	-6.67	1.81	12.06	3.27	3.05
0P00000IYO	Switzerland Equity		2 Quarterly	Blend	Large	288.11	16.32	2.07	1.79	14.59	3.34	6.93	6.82	1.4	9.66	0.66	7.08
0P00000IYQ	Sector Equity Health		3 Quarterly	Growth	Large	312.82	15.78	3.32	2.96	13.28	2.65	8.74	0.48	3.95	7.42	4.43	9.47
0P00000IYS	Sector Equity Techn		2 Quarterly	Growth	Large	284.72	23.8	3.26	3.46	15.86	1.29	14.8	5.47	8.99	8.78	14.02	10.82
0P00000IYW	Sector Equity Financ		2 Quarterly	Value	Large	250.27	8.44	0.39	0.99		1.4	7.13	-5.12	-8.47		0.42	0.18
0P00000IZ2	Sector Equity Financ		1 Quarterly	Value	Large	240.33	7.91	0.33	0.77		1.86	5.65	-4.72	-10.1		0.55	0.13
0P00000IZ6	Global Large-Cap Bl		5 Quarterly	Growth	Large	380.84	21.62	4.24	3.45	14.57	1.92	9.63	10.88	7.41	8.08	4.42	11.02
0P00000IZA	Sector Equity Techn		2	Growth	Large	215.9	22.29	2.75	1.75	8.64	1.49	9.39	12.64	5.38	9.77	6.1	5.43
0P00000IZC	Germany Small/Mid-		3	Growth	Medium	198.02	19.71	1.67	0.94	11.75	1.27	8.08	3.31	3.33	1.74	5.27	5.46
0P00000IZE	Germany Equity		4	Blend	Large	296.65	13.11	1.31	0.85	6.5	2.92	8.25	-8.81	-0.65	3	8.04	1.81
0P00000IZG	Sector Equity Comm		2 Quarterly	Blend	Large	252.8	12.9	1.27	0.81	2.64	4.22	12.49	-1.45	-1.49	6.52	-3.36	2.45
0P00000IZK	Europe Large-Cap Bl		3 Quarterly	Blend	Large	333.64	14.26	1.8	1.52	10.03	3.48	8.98	-3.93	2.13	9.46	3.39	5.75
0P00000K38	Finland Equity		4	Blend	Large	217.48	14.67	1.66	1.34	9.11	3.19	8.95	-5.11	3.42	1.95	0.55	6.94
0P00000KG0	Eurozone Large-Cap		3 Annually	Blend	Large	312.2	14.63	1.47	1.16	9.09	2.76	11.04	-2.63	2.51	7.13	5.47	3.8
0P00000KG2	US Large-Cap Blend		4 Annually	Blend	Large	293.68	19.71	3.13	2.31	13.33	1.91	9.36	10.8	7.01	7.15	6.57	8.28
0P00000KG8	UK Large-Cap Equity		3 Annually	Value	Large	266.9	12.52	1.31	1.02	7.47	4.1	6.62	-6.68	1.83	12.32	3.27	3.11
0P00000MRI	Eurozone Large-Cap		3 Annually	Blend	Large	312.85	14.92	1.54	1.2	9.18	2.71	11.04	-2.63	2.51	7.13	5.47	4.03
0P00000MRP	France Equity		4 Annually	Blend	Large	299.2	14.62	1.35	1.07	9.32	2.21	12.94	4.06	0.55	8.83	3.84	3.69
0P00000MSR	Eurozone Large-Cap		3 Quarterly	Blend	Large	312.43	14.68	1.48	1.17	9.11	2.75	11.04	-2.63	2.51	7.13	5.47	3.67
0P00000RO1	Europe Large-Cap Bl		3 Quarterly	Blend	Large	333.63	14.25	1.8	1.52	10.03	3.48	8.98	-3.93	2.13	9.46	3.39	5.75
0P00000RO2	Eurozone Large-Cap		3 Quarterly	Blend	Large	312.43	14.68	1.48	1.17	9.11	2.75	11.05	-2.63	2.51	7.13	5.47	3.87
0P00000T5O	Italy Equity		3 Annually	Value	Large	237.39	11.77	0.91	0.66	6.06	2.85	7.75	-1.58	0.22	1.76	1.23	2.5
0P00000YK2	UK Mid-Cap Equity		3 Quarterly	Blend	Medium	129.25	14.27	1.39	0.84	7.01	2.74	15.74	-9.43	2.56	5.1	4.23	4.93
0P00000ZF4	Japan Large-Cap Eq		3 Annually	Blend	Large	286.46	17.86	1.2	0.88	8.44	2.74	9.43	-4.08	3.55	2.91	4.92	4.74
0P00000ZH1	Sector Equity Indust		1 Quarterly	Value	Large	249.32	10.35	1.33	0.73	6.43	4.85	11.58	5.86	4.84	11.12	2.18	5.59
0P000015NM	Eurozone Mid-Cap E		3 Quarterly	Blend	Medium	195.84	13.11	1.08	0.7	6.46	2.08	8.58	1.85	0.94	7.01	5.16	3.52
0P000015NT	Eurozone Mid-Cap E		4 Quarterly	Blend	Medium	152.74	14.96	1.01	0.8	7.12	2.72	8.1	3.57	1.85	-0.66	3.24	3.38
0P000015O4	US Large-Cap Growt		5 Quarterly	Growth	Large	351.54	27.19	6.51	4.27	16.55	0.89	13.02	10.01	11.54	10.15	8.91	11.51
0P000015Z4	Sector Equity Consu		1 Quarterly	Value	Large	264.41	9.85	0.79	0.38	4.43	0.72	13.84	-8.7	2.38	3.8	5.81	1.14
0P000015Z5	Sector Equity Financ		3 Quarterly	Value	Large	245.23	7.81	0.83	0.52		5.05	4.34	4.39	-0.21	-25.88	5.54	0.7
0P000015Z6	Sector Equity Indust		4 Quarterly	Growth	Large	278.77	21.27	2.01	1.64	9.08	3.19	9.14	-14.44	1.58	-6.37	23.76	3.63
0P000015Z7	Sector Equity Indust		3 Quarterly	Blend	Large	230.05	15.74	1.87	0.83	7.79	2.28	10.12	14.28	4.91	18.84	3.84	4.78

Problem Guideline Topik 04

1. Lakukan eksplorasi sederhana pada data dan ceritakan insight yang Anda dapatkan dari data tersebut! (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Tunjukkan 5 besar Investment Management berdasarkan dana kelolaan (fund_size)
 - b. Tunjukkan perbandingan untuk tiap sektor (sector_*)
 - c. Apakah terdapat hubungan antara management_fees dengan pertumbuhan return investasi?
 - d. Identifikasi perbedaan antar equity_style

Mahasiswa perlu menambahkan **dua poin eksplorasi sendiri**

2. Buatlah beberapa model yang mungkin akan diperlukan, di antaranya: (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Buatlah model untuk mengklasifikasi "rating"
 - b. Buatlah model untuk memprediksi "long_term_projected_earnings_growth"
 - c. Lakukan clustering pada dataset, dan berikan analisis dari jenis-jenis manajer investasi yang terdapat dalam suatu cluster!

05

Valorant Pro Matches



Deskripsi Topik 05

Deskripsi Dataset (Matches.csv)

Berisi informasi tentang tim yang berpartisipasi dalam sebuah turnamen. Pertandingan umumnya diadakan dengan aturan "best of 3," di mana tim yang memenangkan babak terbanyak akan menjadi pemenang pertandingan.

Jumlah Kolom x Baris

12 kolom x 6401 baris

Deskripsi Kolom

1. MatchID : Id pertandingan
2. Date : Tanggal pertandingan
3. Patch : Versi Valorant saat pertandingan dilakukan
4. EventID : Id turnamen
5. EventName : Nama turnamen
6. EventStage : Tahap turnamen
7. Team1ID : Id tim 1
8. Team2ID : Id tim 2
9. Team1 : Nama tim 1
10. Team2 : Nama tim 2
11. Team1_MapScore : Banyak babak yang dimenangkan tim 1
12. Team2_MapScore : Banyak babak yang dimenangkan tim 2

Deskripsi Topik 05

Gambaran Dataset

index	MatchID	Date	Patch	EventID	EventName	EventStage	Team1ID	Team2ID	Team1	Team2	Team1_MapScore	Tei
0	62393	2022-01-08 15:30:00	Patch 3.12	826	Nerd Street Gamers Winter Championship - Regional LAN Philadelphia	Group Stage: Decider (A)	6903	6020	Booster Seat Gaming	Pho Real	2	
1	62403	2022-01-08 15:30:00	Patch 3.12	826	Nerd Street Gamers Winter Championship - Regional LAN Philadelphia	Group Stage: Decider (C)	7046	7047	Bjor's Kittens	Mugiwara	2	
2	62391	2022-01-08 12:30:00	Patch 3.12	826	Nerd Street Gamers Winter Championship - Regional LAN Philadelphia	Group Stage: Winner's (A)	6461	6903	Akrew	Booster Seat Gaming	2	
		2022-	Patch		Nerd Street Gamers Winter	Group						

Deskripsi Topik 05

Deskripsi Dataset (Games.csv)	Mendokumentasikan detail babak yang dimainkan oleh kedua tim yang bersaing dalam turnamen.
Jumlah Kolom x Baris	18 kolom x 12961 baris
Deskripsi Kolom	

1. GameID : Id pertandingan
2. MatchID : Id babak
3. Map : Peta yang digunakan
4. Team1ID : Id tim 1
5. Team2ID : Id tim 2
6. Team1 : Nama tim 1
7. Team2 : Nama tim 2
8. Winner : Tim yang menjadi pemenang
9. Team1_Eco : Banyak ronde eco yang dilakukan tim 1
10. Team1_SemiEco : Banyak ronde semi eco yang dilakukan tim 1
11. Team1_SemiBuy : Banyak ronde semi buy yang dilakukan tim 1
12. Team1_FullBuy : Banyak ronde full buy yang dilakukan tim 1
13. Team1_TotalRounds : Banyak ronde yang dimenangkan tim 1
14. Team2_Eco : Banyak ronde eco yang dilakukan tim 2
15. Team2_SemiEco : Banyak ronde semi eco yang dilakukan tim 2
16. Team2_SemiBuy : Banyak ronde semi buy yang dilakukan tim 2
17. Team2_FullBuy : Banyak ronde full buy yang dilakukan tim 2
18. Team2_TotalRounds : Banyak ronde yang dimenangkan tim 2

Deskripsi Topik 05

Gambaran Dataset

index	GameID	MatchID	Map	Team1ID	Team2ID	Team1	Team2	Winner	Team1_Eco	Team1_SemiEco	Team1_SemiBuy	Team1
0	60894	62393	Breeze	6903	6020	Booster Seat Gaming	Pho Real	1	2.0	0.0	5.0	
1	60895	62393	Bind	6903	6020	Booster Seat Gaming	Pho Real	2	3.0	3.0	5.0	
2	60896	62393	Haven	6903	6020	Booster Seat Gaming	Pho Real	1	2.0	2.0	5.0	
3	60924	62403	Icebox	7046	7047	Bjor's Kittens	Mugiwara	1	2.0	2.0	4.0	
4	60925	62403	Haven	7046	7047	Bjor's Kittens	Mugiwara	1	4.0	2.0	3.0	
5	60888	62391	Breeze	6461	6903	Akrew	Booster Seat Gaming	2	3.0	2.0	2.0	
6	60889	62391	Bind	6461	6903	Akrew	Booster Seat Gaming	1	3.0	0.0	4.0	
7	60890	62391	Haven	6461	6903	Akrew	Booster Seat Gaming	1	2.0	1.0	4.0	
8	60903	62396	Breeze	6164	7043	Radiance	sameROFLMAO	1	2.0	0.0	4.0	
9	60904	62396	Icebox	6164	7043	Radiance	sameROFLMAO	1	2.0	0.0	5.0	
10	60921	62402	Bind	7047	3545	Mugiwara	CLG Red	1	3.0	1.0	3.0	

Deskripsi Topik 05

Deskripsi Dataset (Scores.csv)	Merupakan catatan akhir dari status para pemain profesional yang bertanding dalam babak tersebut
Jumlah Kolom x Baris	28 kolom x 128779 baris
Deskripsi Kolom	

1. GameID : Id babak
2. PlayerID : Id pemain
3. PlayerName : Username pemain
4. TeamAbbreviation : Singkatan nama tim
5. Agent : Agent/karakter yang digunakan
6. ACS : Skor rata-rata dari pemain
7. Kills : Jumlah musuh yang dikalahkan
8. Deaths : Total pemain berhasil dieliminasi
9. Assists : Jumlah kontribusi pemain dalam membantu rekannya untuk mengalahkan musuh
10. PlusMinus : Selisih antara jumlah mengalahkan musuh (Kills) dan jumlah tereliminasi (Deaths)
11. KAST_Percent : Besar kontribusi pemain dalam pertandingan
12. ADR : Rata-rata serangan yang diberikan oleh pemain ke pihak musuh
13. HS_Percent : Persentase serangan yang diberikan mengenai kepala musuh
14. FirstKills : Total pemain berhasil melakukan eliminasi pertama kali pada suatu ronde
15. FirstDeaths : Total pemain dieliminasi pertama kali pada suatu ronde
16. FKFD_PlusMinus : Selisih antara jumlah First Kills dan First Deaths
17. Num_2Ks, Num_3ks, Num_4ks, Num_5ks : Jumlah pemain berhasil mengeliminasi X musuh dalam suatu ronde
18. OnevOne, OnevTwo, OnevThree, OnevFour, OnevFive : Jumlah pemain berhasil mengalahkan X musuh pada suatu timeframe
19. Econ : Besar serangan yang dihasilkan untuk setiap 1000 creds (mata uang pada permainan) yang digunakan untuk membeli perlengkapan
20. Plants : Jumlah pemain berhasil menanam spike (bom) pada suatu ronde
21. Defuses : Jumlah pemain berhasil menjinakkan spike (bom) pada suatu ronde

Deskripsi Topik 05

Gambaran Dataset

index	GameID	PlayerID	PlayerName	TeamAbbreviation	Agent	ACS	Kills	Deaths	Assists	PlusMinus	KAST_Percent	ADR
0	60894	8419.0	Reduxx	Boos	jett	313.0	24.0	10.0	3.0	14.0	0.65	195.0
1	60894	466.0	ChumZ	Boos	chamber	227.0	16.0	10.0	7.0	6.0	0.9	161.0
2	60894	3712.0	diaamond	Boos	sova	226.0	17.0	9.0	8.0	8.0	0.85	148.0
3	60894	5099.0	Boltzy	Boos	viper	218.0	17.0	12.0	2.0	5.0	0.7	141.0
4	60894	3983.0	Virtyy	Boos	skye	80.0	5.0	13.0	3.0	-8.0	0.7	55.0
5	60894	16757.0	Cohburg	Pho	astra	229.0	15.0	16.0	5.0	-1.0	0.55	155.0
6	60894	19293.0	evaD	Pho	sova	203.0	12.0	14.0	6.0	-2.0	0.65	151.0

HS_Percent	FirstKills	FirstDeaths	FKFD_PlusMinus	Num_2Ks	Num_3Ks	Num_4Ks	Num_5Ks
31.0	4.0	4.0	0.0	2.0	2.0	2.0	0.0
16.0	1.0	1.0	0.0	3.0	1.0	0.0	0.0
27.0	3.0	0.0	3.0	1.0	2.0	0.0	0.0
19.0	3.0	0.0	3.0	3.0	1.0	0.0	0.0
22.0	3.0	1.0	2.0	1.0	0.0	0.0	0.0
31.0	1.0	2.0	-1.0	3.0	2.0	0.0	0.0
22.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0

Problem Guideline Topik 05 (1)

1. Lakukan eksplorasi sederhana pada data dan ceritakan insight yang Anda dapatkan dari data tersebut! (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Bagaimana distribusi Agent yang digunakan oleh pemain untuk setiap Patch 1.X, 2.X, dan 3.X? Apakah terdapat Agent yang mengalami penurunan pengguna pada patch yang berbeda?
 - b. Bagaimana persebaran nilai ACS dari setiap Agent yang digunakan? Apakah pemilihan map yang berbeda menghasilkan persebaran nilai ACS yang berbeda juga?
 - c. Siapa tim yang paling banyak memenangkan pertandingan grand final dari sebuah Event?
 - d. Komposisi tim seperti apakah yang paling banyak memenangkan pertandingan?
 - e. Mahasiswa perlu menambahkan **dua poin eksplorasi sendiri**

Problem Guideline Topik 05 (2)

2. Buatlah beberapa model yang mungkin akan diperlukan, di antaranya: (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
- a. Lakukan klasifikasi untuk menentukan Agent yang digunakan oleh setiap pemain pada setiap babak!
 - b. Buatlah model untuk memprediksi nilai rata-rata ACS dari masing-masing tim untuk sebuah pertandingan!
 - c. Lakukan clustering pada dataset, dan berikan analisis dan insight yang kalian dapatkan dari cluster yang dibuat.

06

Spesifikasi Motor



Deskripsi Topik 06

Deskripsi Dataset (Matches.csv)

Berisi informasi tentang tim yang berpartisipasi dalam sebuah turnamen. Pertandingan umumnya diadakan dengan aturan "best of 3," di mana tim yang memenangkan babak terbanyak akan menjadi pemenang pertandingan.

Jumlah Kolom x Baris

25 kolom x 29966 baris

Deskripsi Kolom

1. Model : Nama brand dan model dari motor
2. Tahun : Tahun motor diproduksi
3. Rating : Rating motor
4. Power : Output power maksimum dalam Horsepower (HP)
5. Kapasitas Bahan Bakar : Kapasitas Bahan Bakar Maksimum dalam Liter (Liter)
6. Jarak Sumbu Roda : Jarak antara ban depan dan ban belakang dalam milimeter (mm)
7. Ketinggian Tempat Duduk : Ketinggian tempat duduk dari bawah tempat duduk hingga dasar dalam milimeter (mm)
8. Bore : Diameter silinder piston dalam milimeter (mm)
9. Stroke : Jarak silinder piston bergerak dalam milimeter (mm)
10. Jenis Mesin : Jenis mesin yang digunakan oleh motor
11. Jenis Bahan Bakar : Jenis bahan bakar yang digunakan oleh motor
12. Jenis Sistem Pendingin : Jenis sistem pendingin yang digunakan oleh motor
13. Gearbox : Banyaknya gear dalam gearbox.
14. Jenis Transmisi : Jenis transmisi yang digunakan oleh motor
15. Suspensi Depan : Jenis dan konfigurasi suspensi bagian depan
16. Suspensi Belakang : Jenis dan konfigurasi suspensi bagian belakang

Deskripsi Topik 06

Deskripsi Kolom

- 17. Kompresi : Rasio kompresi yang digunakan oleh motor
- 18. Ban Depan : Jenis dan ukuran yang digunakan oleh ban depan
- 19. Ban Belakang: Jenis dan ukuran yang digunakan oleh ban belakang
- 20. Rem Depan : Jenis rem depan yang digunakan
- 21. Rem Belakang: Jenis rem belakang yang digunakan
- 22. Wana : Warna yang digunakan pada motor
- 23. Starter : Jenis starter yang digunakan untuk menyalakan motor
- 24. Kapasitas Mesin : Ukuran kapasitas mesin dalam centimeter cubic (cc)
- 25. Kategori: Kategori motor

Deskripsi Topik 06

Gambaran Dataset

Model	Tahun	Rating	Power	Kapasita	Jarak Sui	Ketinggia	Bore	Stroke	Jenis Me	Jenis Bak	Jenis Sisi	Gearbox	Jenis Tra	Suspens	Suspens	Kompres	Ban Dep.	Ban Bela	Rem Dep	Rem Bela	Warna	Starter	Kapasita	Kategori
Adiva Bir	2012	2.2	9.4	5	1400	800	54	66	Single cy	Carbureti	Air	6-speed	Chain Å	Telescop	Monoshc	9.5:1	3.50-10	3.50-10	Expandir	Expandir	Red/white	Electric	124.6	Classic
Yamaha	1997	3.7	48	17	1400	690	54	66	Single cy	Carbureti	Liquid	5-speed	Chain Å	Telescop	Monoshc	9.5:1	120/70-Z	180/55-Z	Single dis	Single dis	Black	Electric	660	Classic
Modenas	2011	3	27	4.3	1400	800	50	49.5	Single cy	Carbureti	Air	4-speed	Belt Å (fir	Telescop	Trailing st	9.0:1	120/70-Z	180/55-Z	Expandir	Expandir	Blue, bla	Kick	97.2	Classic
Triumph	1956	3.5	27	17	1400	800	54	66	Twin, fou	Carbureti	Air	6-speed	Chain Å	Telescop	Rigid-opt	9.5:1	120/70-Z	180/55-Z	Expandir	Expandir	Black	Electric	498	Classic
Sym Mio	2007	2.8	27	4.8	1220	800	54	66	Single cy	Carbureti	Air	6-speed	Chain Å	Telescop	Monoshc	9.5:1	120/70-Z	180/55-Z	Single dis	Expandir	Black, Gr	Electric 8	100	Classic
Honda V	2007	3.7	27	18.16	1662	697	89.5	104.3	Twin, fou	Carbureti	Liquid	5-speed	Shaft driv	Telescop	Dual shor	9.2:1	110/90-1	170/80-1	Single dis	Single dis	Black, D	Electric	1312	Cruiser
Yamaha	2003	3	21.1	11	1530	800	54	66	V2, four-	Carbureti	Air	6-speed	Chain Å	Telescop	Monoshc	9.5:1	120/70-Z	180/55-Z	Single dis	Single dis	Black	Electric	249	Cruiser
Gibbs Qu	2013	3.5	27	17	1800	800	54	66	In-line fo	Carbureti	Air	6-speed	Chain Å	Indepenc	Indepenc	9.5:1	120/70-Z	180/55-Z	Single dis	Single dis	Red, blue	Electric	1300	Offroad
Harley-D	2003	4	27	19.7	1695	720	95.3	101.6	V2, four-	Carbureti	Air	6-speed	Belt Å (fir	Telescop	Monoshc	8.8:1	120/70-Z	180/55-Z	Single dis	Single dis	Vivid blac	Electric	1449	Cruiser
Bajaj Pul	2010	3.3	14.1	15	1320	800	54	66	Single cy	Carbureti	Air	6-speed	Chain Å	Telescop	Triple rat	9.5:1	90/90-17	120/80-1	Single dis	Expandir	Silver, bl	Electric	149	Sport
MV Agus	2008	3.4	171.7	21	1408	800	76	55	In-line fo	Injection	Liquid	6-speed	Chain Å	Upside-c	Progress	13.0:1	120/70-Z	190/55-Z	Double d	Single dis	Black/rec	Electric	998	Sport
MZ RT 12	2005	3.7	11	12.5	1355	770	54	66	Single cy	Carbureti	Air	6-speed	Chain Å	Telescop	Rear mor	9.5:1	110/70-1	130/70-1	Single dis	Single dis	Black	Electric	124	Sport
Aprilia SF	2011	3	27	7	1290	820	40	39.3	Single cy	Carbureti	Liquid	Automati	Belt Å (fir	Hydraulic	Engine ut	12.1:1	130/60-1	130/60-1	Single dis	Single dis	Blue,	Electric	49	Classic
Lem Rec	2007	3.5	3.2	2.49	1080	729	1.5	0	Single cy	Carbureti	Air	1-speed	Chain Å	Mechani	Hydraulic	9.5:1	2.00-14	3.00-12	Single dis	Expandir	Multicolo	Electric 8	49.8	Offroad
Ducati 9C	1990	3.3	73	11.5	1400	800	92	68	V2, four-	Carbureti	Air	6-speed	Chain Å	Telescop	Monoshc	9.2:1	120/70-1	170/60-1	Dual disc	Single dis	Black	Electric	904	Sport
TMRacir	2007	3.5	27	8	1400	800	54	54.5	Single cy	Carbureti	Liquid	5-speed	Chain Å	Marzocol	Sachs B	16.7:1	90/90-21	100/90-1	Single dis	Single dis	Black	Electric	123.5	Offroad
Arlen Nes	2013	3.5	112	18.93	1400	800	101.6	114.3	V2, four-	Carbureti	Air	6-speed	Belt Å (fir	63.5mm i	Monoshc	10.1:1	3.50-23	10.5-18	Single dis	Single dis	Custom	Electric	1851.9	Cruiser
Aprilia Tu	2007	3.6	139	22	1410	810	97	67.5	V2, four-	Injection.	Oil & air	6-speed	Chain Å	Å-hlins R	Å-hlins R	11.8:1	120/70-Z	180/55-Z	Double d	Single dis	Lion red	Electric	997.6	Sport
Suzuki G	2008	3.9	99.2	22	1520	775	54	66	In-line fo	Injection	Oil & air	6-speed	Chain Å	Telescop	Monoshc	9.5:1	120/70-Z	190/50-Z	Double d	Single dis	Black	Electric	1401	Classic
Highland	2007	3.5	83.8	20	1530	910	95	66	V2, four-	Injection.	Liquid	6-speed	Chain Å	50 mm M.	Å-hlins, C	9.5:1	90/90-21	150/70-1	Single dis	Single dis	Yellow, si	Electric	936	Offroad
Lem LX 2	2005	3.5	27	2	1400	572	39	41.7	Single cy	Carbureti	Air	Automati	Chain Å	Telescop	Monoshc	9.5:1	120/70-Z	180/55-Z	Expandir	Expandir	Black	Kick	49.8	Offroad
Headbar	2011	3.5	27	18	1665	800	54	66	V2, four-	Carbureti	Air	5-speed	Chain Å	41mm tel	Monoshc	9.5:1	130/80-1	180/55-1	Single dis	Single dis	Black	Electric	1638.8	Cruiser
Honda V	1998	4	110	17	1400	810	54	66	V2, four-	Carbureti	Liquid	6-speed	Chain Å	Telescop	Monoshc	9.5:1	120/70-Z	180/55-Z	Dual disc	Single dis	Black	Electric	996	Sport
Norton 4	1928	3.5	27	17	1400	800	54	66	Single cy	Carbureti	Air	6-speed	Chain Å	Telescop	Monoshc	9.5:1	120/70-Z	180/55-Z	Expandir	Expandir	Black	Electric	490	Classic
Rokon R	2019	3.5	7	10	1295	813	54	66	Single cy	Carbureti	Liquid	3-speed	Chain Å	Telescop	Monoshc	9.5:1	120/70-Z	180/55-Z	Single dis	Single dis	Black, oli	Electric 8	208	Offroad
Chang-J	2007	3.5	7.8	5	1232	800	52.4	57.8	Single cy	Carbureti	Air	6-speed	Belt Å (fir	Telescop	Monoshc	9.2:1	3.5-10	3.5-10	Single dis	Expandir	Black	Electric 8	125	Classic
Malaguti	2009	3.5	27	9.5	1290	830	54	66	Single cy	Carbureti	Air	Automati	Belt Å (fir	Telescop	Monoshc	9.5:1	130/60-1	130/60-1	Single dis	Expandir	Black	Electric 8	49	Classic
ZÅ/inda	1955	3.5	28	17	1400	800	75	67.6	Two cylin	Carbureti	Air	4-speed	Shaft driv	Telescop	Monoshc	9.5:1	3.50-19	3.50-19	Expandir	Expandir	Black	Kick	592	Classic
Harley-D	2004	3.6	27	17	1400	800	54	66	V2, four-	Carbureti	Air	5-speed	Belt Å (fir	Telescop	Monoshc	9.5:1	120/70-Z	180/55-Z	Single dis	Single dis	Black	Electric	1556.9	Cruiser
Suzuki Bi	2022	3.5	8.7	5.5	1265	780	52.5	57.4	Single cy	Injection	Air	Automati	Belt Å (fir	Telescop	Swingarn	11.6:1	90/90-12	90/100-1	Single dis	Expandir	Black, w	Electric 8	124	Classic
Honda XI	1983	3	10	7	1400	800	63	57.8	Single cy	Carbureti	Air	5-speed	Chain Å	Telescop	Monoshc	9.2:1	2.75-21	4.10-18	Expandir	Expandir	Black	Electric	180	Offroad
Derbi GP	2004	3.9	27	7	1280	800	40	39.3	Single cy	Carbureti	Liquid	6-speed	Chain Å	Hydraulic	Linear sir	10.5:1	120/70-1	130/60-1	Single dis	Single dis	Black	Electric	49.4	Classic

Problem Guideline Topik 06

1. Lakukan eksplorasi sederhana pada data dan ceritakan insight yang Anda dapatkan dari data tersebut! (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Tunjukkan brand yang memiliki motor rating terbagus?
 - b. Tunjukkan tren power/kekuatan berdasarkan keluaran tahun motor?
 - c. Kategori motor apa yang paling banyak diproduksi di setiap tahunnya?
 - d. Apa model motor scooter dengan rating terbaik yang diproduksi oleh lima brand teratas yang memproduksi motor paling banyak?

Mahasiswa perlu menambahkan **dua poin eksplorasi sendiri**

2. Buatlah beberapa model yang mungkin akan diperlukan, di antaranya: (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Buatlah model untuk mengklasifikasi Kategori suatu motor.
 - b. Buatlah model untuk memprediksi nilai kapasitas mesin dari suatu motor.
 - c. Lakukan clustering pada dataset, dan berikan analisis dari clustering yang dibuat!

07

Telco Customer Behavior (Marketing)



Deskripsi Topik 07

Deskripsi Dataset	Dataset ini menunjukkan behavior customer dan marketing mix yang didapatkannya dari sebuah perusahaan telco.
Jumlah Kolom x Baris	25 kolom x 26,277 baris
Deskripsi Kolom	

- 'CustomerID': ID of customer
- 'Gender': jenis kelamin
- 'Senior Citizen': flag apakah customer berusia lanjut
- 'Partner': flag apakah customer mempunyai pasangan
- 'Dependents': flag apakah customer mempunyai dependents/tanggungan
- 'Tenure': jumlah bulan customer sudah menjadi pelanggan
- 'Phone Service': flag apakah customer menggunakan layanan telpon
- 'Multiple Lines': flag apakah customer menggunakan layanan multi-line
- 'Internet Service': flag apakah customer menggunakan layanan internet
- 'Online Security': flag apakah customer menggunakan layanan keamanan online

Deskripsi Topik 07

Deskripsi Kolom

- 'Online Backup': flag apakah customer menggunakan layanan backup data online
- 'Device Protection': flag apakah customer menggunakan layanan proteksi gawai
- 'Tech Support': flag apakah customer pernah menggunakan pusat bantuan teknis
- 'Streaming TV': flag apakah customer menggunakan layanan streaming TV
- 'Streaming Movies': flag apakah customer menggunakan layanan streaming film
- 'Contract': jenis durasi kontrak yang dipilih customer
- 'Paperless Billing': flag apakah user menggunakan paperless billing
- 'Payment Method': metode pembayaran yang digunakan customer
- 'Monthly Charges': besar tagihan bulanan yang dibayarkan customer
- 'Churn': flag apakah customer memutuskan berhenti menggunakan layanan telco dari company ini
- 'Number of Dependents': jumlah tanggungan
- 'Mobile App Usage': flag apakah customer menggunakan mobile app yang disediakan company (seperti aplikasi MyTelkomsel)
- 'Marketing Number of Cold Call': jumlah cold call yang pernah diterima oleh customer
- 'Marketing Number of Email': jumlah email marketing yang pernah diterima oleh customer
- 'Marketing Number of SMS': jumlah SMS marketing yang pernah diterima oleh customer

Deskripsi Topik 07

Gambaran Dataset

index	CustomerID	Gender	Senior Citizen	Partner	Dependents	Tenure	Phone Service	Multiple Lines	Internet Service	Online Security	Online Backup	Device Protection	Tech Support	Streaming TV	Streaming Movies	Contract	Paperless Billing	Payment Method
0	7590-VHVEG	0	0.0	Yes	No	NaN	No	No phone service	DSL	No	Yes	No	No	No	No	Month-to-month	NaN	Electronic check
1	5575-GNVDE	1	0.0	NaN	No	NaN	Yes	No	DSL	Yes	No	Yes	No	NaN	No	One year	No	Mailed check
2	3668-QPYBK	1	0.0	No	No	2.0	NaN	No	DSL	Yes	Yes	No	No	No	NaN	Month-to-month	Yes	Mailed check
3	9237-HQITU	0	NaN	No	No	2.0	NaN	No	Fiber optic	No	No	No	No	No	No	Month-to-month	NaN	Electronic check
4	9305-CDSKC	0	0.0	No	NaN	NaN	Yes	Yes	Fiber optic	No	No	Yes	No	Yes	Yes	Month-to-month	NaN	Electronic check
5	1452-KIOVK	1	0.0	No	NaN	22.0	Yes	Yes	Fiber optic	No	Yes	No	No	Yes	No	Month-to-month	NaN	Credit card (automatic)
6	6713-OKOMC	0	0.0	No	No	10.0	No	No phone service	DSL	Yes	NaN	No	No	No	NaN	Month-to-month	No	Mailed check
7	7892-POOKP	0	NaN	NaN	NaN	28.0	NaN	Yes	Fiber optic	No	No	Yes	Yes	Yes	Yes	Month-to-month	Yes	Electronic check
8	9763-GRSKD	1	0.0	Yes	NaN	13.0	Yes	No	DSL	Yes	NaN	No	No	No	No	Month-to-month	NaN	Mailed check
9	7469-LKBCI	1	0.0	NaN	NaN	16.0	Yes	No	No	No internet service	No internet service	No internet service	NaN	No internet service	No internet service	Two year	NaN	Credit card (automatic)
10	0280-XJGEX	1	0.0	NaN	No	49.0	NaN	Yes	Fiber optic	No	Yes	Yes	No	Yes	Yes	Month-to-month	Yes	Bank transfer (automatic)
11	5129-JLPIS	1	0.0	No	No	NaN	Yes	No	Fiber optic	Yes	No	Yes	Yes	Yes	Yes	Month-to-month	Yes	Electronic check
12	8191-XWSZG	0	0.0	No	NaN	NaN	Yes	No	No	No internet service	No internet service	No internet service	NaN	No internet service	No internet service	One year	No	Mailed check
13	9959-WOFKT	1	0.0	No	Yes	NaN	Yes	Yes	NaN	Yes	No	Yes	No	Yes	Yes	Two year	No	Bank transfer (automatic)
14	4190-MFLUW	0	0.0	Yes	Yes	10.0	Yes	No	DSL	No	No	Yes	Yes	No	No	Month-to-month	NaN	NaN
15	4183-MYFRB	0	0.0	No	No	21.0	Yes	No	Fiber optic	No	Yes	Yes	No	No	NaN	Month-to-month	NaN	Electronic check
16	8779-QRDMV	1	1.0	No	NaN	1.0	No	No phone service	DSL	No	No	Yes	No	NaN	Yes	Month-to-month	Yes	Electronic check
17	1680-VDCWW	1	0.0	Yes	No	NaN	Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	One year	NaN	NaN
18	1066-JKSGK	1	0.0	No	No	NaN	Yes	No	No	No internet service	No internet service	No internet service	No internet service	NaN	No internet service	Month-to-month	NaN	Mailed check
19	3638-WEABW	0	NaN	Yes	No	58.0	NaN	Yes	DSL	No	NaN	No	Yes	No	No	Two year	Yes	Credit card (automatic)

Problem Guideline Topik 07

1. Lakukan eksplorasi sederhana pada data dan ceritakan insight yang Anda dapatkan dari data tersebut! (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Periksa payment method apa yang banyak digunakan berdasarkan jenis kelamin dan umur.
 - b. Untuk customer 25-30 tahun, layanan apa saja yang akan kamu rekomendasikan?
 - c. Eksplor bagaimana karakteristik orang-orang yang tidak menggunakan layanan telpon?
 - d. Strategi media marketing mix apa yang sebaiknya diberikan ke customer untuk meningkatkan pembeliannya dan tidak churn?

Mahasiswa perlu menambahkan **dua poin eksplorasi sendiri**

2. Buatlah beberapa model yang mungkin akan diperlukan, di antaranya: (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Buatlah model untuk mengklasifikasi 'contract'. Atribut apa saja yang perlu diperhatikan company ketika ingin mengirimkan penawaran 'contract'?
 - b. Buatlah model untuk memprediksi 'monthly charges' customer. Bagaimana hasil prediksi anda dapat membantu perusahaan mengambil keputusan?
 - c. Lakukan analisis cluster yang dapat terbentuk pada data yang tersedia. Deskripsikan karakteristik masing-masing cluster

08

Ultimate Fighter Championship (UFC)



Deskripsi Topik 08

Deskripsi Dataset	Dataset ini berisi daftar setiap pertandingan UFC dalam sejarah organisasi tersebut. Setiap baris berisi informasi tentang kedua petinju, detail pertandingan, dan pemenangnya.
Jumlah Kolom x Baris	144 kolom x 6.012 baris

Deskripsi Topik 08

Deskripsi Kolom

- **R_** and **B_** : Awalan ini mengindikasikan statistik petinju sudut merah (red corner) dan sudut biru (blue corner) masing-masing.
- **_opp_** : Kolom ini berisi rata-rata kerusakan yang diterima oleh petinju dari lawannya.
- **KD** : Jumlah knockdown.
- **SIG_STR** : Jumlah pukulan signifikan 'terkena dari yang dicoba'.
- **SIG_STR_pct** : Persentase pukulan signifikan yang berhasil.
- **TOTAL_STR** : Jumlah pukulan total 'terkena dari yang dicoba'.
- **TD** : Jumlah takedown.
- **TD_pct** : Persentase takedown yang berhasil.
- **SUB_ATT** : Jumlah upaya submission (penyerahan lawan).
- **REV** : Jumlah Reversals (pembalikan posisi).
- **HEAD** : Jumlah pukulan signifikan ke kepala 'terkena dari yang dicoba'.
- **BODY** : Jumlah pukulan signifikan ke tubuh 'terkena dari yang dicoba'.
- **CLINCH** : Jumlah pukulan signifikan dalam clinch 'terkena dari yang dicoba'.
- **GROUND** : Jumlah pukulan signifikan di atas tanah 'terkena dari yang dicoba'.
- **Referee** : Nama wasit pertandingan.
- **date** : Tanggal pertandingan.
- **location** : Lokasi di mana pertandingan tersebut berlangsung.
- **winner** : Pemenang pertandingan.
- **Stance** : Gaya bertarung petinju (orthodox, southpaw, dll.).
- **Height_cms** : Tinggi petinju dalam sentimeter.
- **Reach_cms** : Rentang tangan petinju (panjang lengan) dalam sentimeter.
- **Weight_lbs** : Berat petinju dalam pound (lbs).

Deskripsi Topik 08

Deskripsi Kolom

- `age`: Umur petinju.
- `title_bout`: Nilai Boolean yang menunjukkan apakah pertandingan ini adalah pertandingan judul atau bukan.
- `weight_class`: Kelas berat pertandingan (Bantamweight, heavyweight, Women's flyweight, dll.).
- `current_lose_streak`: Jumlah kekalahan beruntun saat ini oleh petinju.
- `current_win_streak`: Jumlah kemenangan beruntun saat ini oleh petinju.
- `draw`: Jumlah seri dalam karier UFC petinju.
- `wins`: Jumlah kemenangan dalam karier UFC petinju.
- `losses`: Jumlah kekalahan dalam karier UFC petinju.
- `total_rounds_fought`: Rata-rata total ronde pertarungan yang dilakukan oleh petinju.
- `total_time_fought(seconds)`: Total waktu yang dihabiskan petinju dalam pertarungan, dihitung dalam detik.
- `total_title_bouts`: Total jumlah pertandingan judul yang diikuti oleh petinju.
- `win_by_Decision_Majority`: Jumlah kemenangan dengan keputusan mayoritas oleh juri dalam karier UFC petinju.
- `win_by_Decision_Split`: Jumlah kemenangan dengan keputusan split oleh juri dalam karier UFC petinju.
- `win_by_Decision_Unanimous`: Jumlah kemenangan dengan keputusan bulat oleh juri dalam karier UFC petinju.
- `win_by_KO/TKO`: Jumlah kemenangan dengan knockout dalam karier UFC petinju.
- `win_by_Submission`: Jumlah kemenangan dengan submission dalam karier UFC petinju.
- `win_by_TKO_Doctor_Stoppage`: Jumlah kemenangan dengan penghentian oleh dokter dalam karier UFC petinju.

Deskripsi Topik 08

Gambaran Dataset

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	R_fighter	B_fighter	Referee	date	location	Winner	title_bout	weight_class	B_avg_KD	B_avg_opp_KD	B_avg_SIG_STF	B_avg_opp_SIG	B_avg_TD_pct	B_avg_opp_TD	B_avg_SUB_AT	B_avg_opp_SUB	B_avg_REV	B_avg_opp_REV
2	Adrian Yanez	Gustavo Lopez	Chris Tognoni	2021-03-20	Las Vegas, Nev	Red	FALSE	Bantamweight	0	0	0.42	0.495	0.33	0.36	0.5	1	0	0
3	Trevin Giles	Roman Dolidze	Herb Dean	2021-03-20	Las Vegas, Nev	Red	FALSE	Middleweight	0.5	0	0.66	0.305	0.3	0.5	1.5	0	0	0
4	Tai Tuivasa	Harry Hunsucker	Herb Dean	2021-03-20	Las Vegas, Nev	Red	FALSE	Heavyweight										
5	Cheyenne Buys	Montserrat Cone	Mark Smith	2021-03-20	Las Vegas, Nev	Blue	FALSE	WomenStrawweight										
6	Marion Reneau	Macy Chasson	Mark Smith	2021-03-20	Las Vegas, Nev	Blue	FALSE	WomenBantamw	0.125	0	0.535625	0.57875	0.185	0.16625	0.125	0.1875	0.25	0
7	Leonardo Santos	Grant Dawson	Chris Tognoni	2021-03-20	Las Vegas, Nev	Blue	FALSE	Lightweight	0	0	0.515	0.47375	0.435	0.25	1.25	0.625	0	0.25
8	Song Kenan	Max Griffin	Keith Peterson	2021-03-20	Las Vegas, Nev	Blue	FALSE	Welterweight	0.046875	0.125	0.4592773437	0.4046875	0.3221875	0.0688671875	0.189453125	0	0.283203125	0
9	Derek Brunson	Kevin Holland	Herb Dean	2021-03-20	Las Vegas, Nev	Red	FALSE	Middleweight	0.15625	0	0.7915820312	0.8505859375	0.2691992188	0.3933984375	0.029296875	0.080078125	0.0234375	0.041015625
10	Montel Jackson	Jesse Strader	Keith Peterson	2021-03-20	Las Vegas, Nev	Red	FALSE	Bantamweight										
11	Bruno Silva	JP Buys	Mark Smith	2021-03-20	Las Vegas, Nev	Red	FALSE	Flyweight										
12	Manel Kape	Matheus Nicolau		2021-03-13	Las Vegas, Nev	Blue	FALSE	Flyweight	1	0.625	0.47375	0.24125	0.28125	0	0.875	0	0	0
13	Misha Cirkunov	Ryan Spann	Mark Smith	2021-03-13	Las Vegas, Nev	Blue	FALSE	LightHeavyweig	1.0625	0	0.473125	0.663125	0.60125	0.105	0.5	0.0625	0.0625	0.5
14	Leon Edwards	Belal Muhammad	Herb Dean	2021-03-13	Las Vegas, Nev	Draw	FALSE	Welterweight	0	0.06494140625	0.4657324219	0.3741699219	0.1764648438	0.063125	0.140625	0.00390625	0	0.005859375
15	Charles Jourdain	Marcelo Rojo	Mark Smith	2021-03-13	Las Vegas, Nev	Red	FALSE	Featherweight										
16	Eryk Anders	Darren Stewart	Herb Dean	2021-03-13	Las Vegas, Nev	Draw	FALSE	Middleweight	0.0234375	0	0.4514463125	0.5031933594	0.3235742188	0.3143945313	0.25	0.015625	0	0.125
17	Angela Hill	Ashley Yoder	Mike Beltran	2021-03-13	Las Vegas, Nev	Red	FALSE	WomenStrawwe	0.015625	0	0.496640625	0.61671875	0.3271875	0.0909375	1.59375	0.1484375	0.1796875	0.640625
18	Jonathan Martin	Davey Grant	Chris Tognoni	2021-03-13	Las Vegas, Nev	Blue	FALSE	Bantamweight	0.5	0.625	0.464375	0.398125	0.48375	0.048125	0.5	0.34375	0.28125	1.0625
19	Dan Ige	Gavin Tucker	Mike Beltran	2021-03-13	Las Vegas, Nev	Red	FALSE	Featherweight	0	0.3125	0.529375	0.4075	0.5	0.160625	1.3125	0.3125	0.375	0.25
20	Nasrat Haqparov	Rafa Garcia	Chris Tognoni	2021-03-13	Las Vegas, Nev	Red	FALSE	Lightweight										
21	Courtney Casey	JJ Aldrich	Mark Smith	2021-03-13	Las Vegas, Nev	Blue	FALSE	WomenFlyweigh	0	0	0.4046875	0.37546875	0.13125	0.2728125	0	0	0	0
22	Matthew Semels	Jason Witt	Chris Tognoni	2021-03-13	Las Vegas, Nev	Red	FALSE	Welterweight	0	0.5	0.52	0.525	0.33	0	0.5	0	0	0
23	Rani Yahya	Ray Rodriguez	Herb Dean	2021-03-13	Las Vegas, Nev	Red	FALSE	Bantamweight	0	0	0.2	0.33	1	0	0	1	0	0
24	Gloria de Paula	Jinh Yu Frey	Mike Beltran	2021-03-13	Las Vegas, Nev	Blue	FALSE	WomenStrawwe	0	0	0.49	0.57	0.125	0.09	0	0.5	0	0
25	Dominick Cruz	Casey Kenney	Jason Herzog	2021-03-06	Las Vegas, Nev	Red	FALSE	Bantamweight	0	0	0.4859375	0.381875	0.216875	0.0765625	0.15625	0	0.125	0.0625
26	Song Yadong	Kyler Phillips	Mark Smith	2021-03-06	Las Vegas, Nev	Blue	FALSE	Bantamweight	0	0	0.465	0.335	0.75	0.165	0.5	0.5	0	0
27	Petr Yan	Aljamain Sterling	Mark Smith	2021-03-06	Las Vegas, Nev	Blue	TRUE	Bantamweight	0	0.03515625	0.5821240234	0.4207873535	0.03421875	0.07447509766	1.128051758	0.000122070312	0.03125	0.00390625
28	Jan Blachowicz	Israel Adesanya	Herb Dean	2021-03-06	Las Vegas, Nev	Red	TRUE	LightHeavyweig	1.05078125	0	0.5284765625	0.42078125	0	0.0095703125	0.1328125	0	0	0.0625
29	Thiago Santos	Aleksandar Raki	Herb Dean	2021-03-06	Las Vegas, Nev	Blue	FALSE	LightHeavyweig	0.71875	0.0625	0.621875	0.4659375	0.0509375	0.004375	0.25	0	0	0
30	Joseph Benavidez	Askar Askarov	Jason Herzog	2021-03-06	Las Vegas, Nev	Blue	FALSE	Flyweight	0	0.25	0.5625	0.4625	0.155	0.5675	0	0	0.75	0.25
31	Amanda Nunes	Megan Anderson	Jason Herzog	2021-03-06	Las Vegas, Nev	Red	TRUE	WomenFeatherw	0.5	0	0.498125	0.415	0.125	0.25875	0.25	0.1875	0.5625	0.25
32	Islam Makhachev	Drew Dober	Chris Tognoni	2021-03-06	Las Vegas, Nev	Red	FALSE	Lightweight	0.388671875	0	0.469420166	0.420960083	0.03159179688	0.3411578369	0.000122070312	0.09838867188	0	0.01953125
33	Tim Elliott	Jordan Espinosa	Mark Smith	2021-03-06	Las Vegas, Nev	Red	FALSE	Flyweight	0	0	0.418125	0.369375	0.185625	0.125	0	0.1875	0	0
34	Luivinha Souza	Amanda Lemos	Jason Herzog	2021-03-06	Las Vegas, Nev	Blue	FALSE	WomenStrawwe	0.5	0	0.645	0.555	1	0	0.25	0	0	0

Problem Guideline Topik 08

- Lakukan eksplorasi sederhana pada data dan ceritakan insight yang Anda dapatkan dari data tersebut! (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Siapa petinju dengan jumlah kemenangan terbanyak? Berapa persentase kemenangannya? (Jumlah kemenangan dapat dihitung menggunakan nilai pada kolom `Winner` tanpa menghitung pertandingan yang seri)
 - b. Apakah umur dari petinju dapat menentukan kemenangan mereka? Buatlah visualisasi untuk melihat jumlah kemenangan pada kolom `Winner` berdasarkan umur petinjunya
 - c. Menampilkan hubungan antara fitur panjang jangkauan lengan petinju dari corner biru (`B_Reach_cms`) dengan fitur lainnya yang dinilai relevan
 - d. Pada tahun berapakah yang memiliki UFC Fight Card* terbanyak dari rentang tahun 1994 – 2021.

*UFC Fight Card adalah daftar pertandingan atau pertarungan yang akan diadakan dalam suatu acara UFC (Ultimate Fighting Championship). Hint : Dilihat dari tanggal terjadinya pertandingan, bukan dari jumlah pertandingan yang ada

Mahasiswa perlu menambahkan **dua poin eksplorasi sendiri**

Problem Guideline Topik 08

- Buatlah beberapa model yang mungkin akan diperlukan, di antaranya: (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Buatlah sebuah model yang dapat memprediksi pemenang (`winner`) dari suatu pertandingan UFC. Anda bebas menggunakan fitur apapun pada dataset!
 - b. Buatlah sebuah model yang dapat memprediksi panjang jangkauan lengan petinju dari corner biru (`B_reach_cms`)!
 - c. Buatlah cluster dari dataset dan jelaskan karakteristik dari cluster-cluster yang terbentuk!

09

Customer Credit Loan



Deskripsi Topik 09

Deskripsi Dataset	Dataset ini berisi daftar pinjaman kredit pelanggan. Setiap baris berisi informasi pinjaman pelanggan dalam berbagai data.
Jumlah Kolom x Baris	19 kolom x 85.296 baris

- 'Loan ID' : ID of loan
- 'Customer ID': ID of customer
- 'Loan Status': jenis status peminjaman customer
- 'Customer Loan Amount': jumlah pinjaman customer
- 'Term': jenis term peminjaman customer
- 'Credit Score': skor credit customer
- 'Annual Income': pendapatan tahunan
- 'Years in current job': lama customer telah bekerja
- 'Home Ownership': jenis kepemilikan rumah customer
- 'Purpose': jenis tujuan peminjaman customer

Deskripsi Topik 09

Deskripsi Kolom

- 'Monthly Debt': banyaknya hutang per bulan
- 'Years of Credit History': credit history customer
- 'Months since last delinquent': lama customer telah melewati tenggat pembayaran
- 'Number of Open Accounts': jumlah akun customer yang dibuka
- 'Number of Credit Problems': jumlah masalah credit yang dimiliki customer
- 'Current Credit Balance': jumlah saldo yang dimiliki oleh customer
- 'Maximum Open Credit': jumlah credit yang dibuka maksimal
- 'Bankruptcies': flag apakah customer bangkrut
- 'Tax Liens': flag apakah customer tersebut terkena penyitaan karena pajak

1a. curr loan amount, pendapatan tahunan
2

Problem Guideline Topik 09

1. Lakukan eksplorasi sederhana pada data dan ceritakan insight yang Anda dapatkan dari data tersebut! (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Fitur apakah yang berkaitan paling tinggi dengan tingkat skor credit dari antara jumlah peminjaman, pendapatan tahunan, hutang bulanan, dan jumlah saldo kredit saat ini?
 - b. Apakah tujuan peminjaman dengan persentase terbanyak berdasarkan masing-masing lama waktu peminjam telah bekerja?
 - c. Berapa rata-rata hutang bulanan dari peminjaman dengan berbagai macam jenis kepemilikan rumah?
 - d. Bagaimana distribusi pendapatan peminjam berdasarkan kepemilikan rumah dan jumlah akun yang akan dibuka?

Mahasiswa perlu menambahkan **dua poin eksplorasi sendiri**

-
2. Buatlah beberapa model yang mungkin akan diperlukan, di antaranya: (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Buatlah model untuk mengklasifikasi 'Loan Status'. Atribut apa saja yang perlu diperhatikan ketika ingin menentukan 'Loan Status' customer?
 - b. Buatlah model untuk memprediksi 'Maximum Open Credit' customer. Bagaimana hasil prediksi anda dapat membantu menentukan credit maksimal yang dapat dibuka oleh customer?
 - c. Lakukan analisis cluster yang dapat terbentuk pada data yang tersedia. Deskripsikan karakteristik masing-masing cluster

10

Meramal Cuaca (MC)



Deskripsi Topik 10

Deskripsi Dataset	Dataset yang diberikan merupakan dataset yang berkaitan dengan cuaca curah hujan yang nilai waktunya dapat berupa barisan. Dataset ini terdiri dari 2 jenis yaitu penelitian informasi di tingkat hari dan di tingkat jam.
Jumlah Kolom x Baris	Harian: 17 kolom x 14730 baris, Jam: 27 kolom x 353520 baris
Deskripsi Kolom	

Deskripsi Topik 10

Data harian:

- `Id` - id cuaca
- `time` - Tanggal pencatatan
- `temperature_2m_max` (°C) - Temperatur udara tertinggi pada ketinggian 2 m di atas permukaan
- `temperature_2m_min` (°C) - Temperatur udara terendah pada ketinggian 2 m di atas permukaan
- `apparent_temperature_max` (°C) - Temperatur semu maksimum yang terasa
- `apparent_temperature_min` (°C) - Temperatur semu minimum yang terasa
- `sunrise (iso8601)` - Waktu matahari terbit pada hari itu dengan format ISO 8601
- `sunset (iso8601)` - Waktu matahari tenggelam pada hari itu dengan format ISO 8601
- `shortwave_radiation_sum` (MJ/m²) - Total radiasi matahari pada hari tersebut
- `rain_sum` (mm) - Jumlah curah hujan pada hari tersebut - Target Regresi

Deskripsi Topik 10

- snowfall - Jumlah salju pada hari tersebut
- precipitation - Jumlah curah hujan harian (termasuk hujan, hujan lebat, dan salju)
- windspeed_10m_max (km/h) - Kecepatan angin maksimum pada ketinggian 10 m
- windgusts_10m_max (km/h) - Kecepatan angin minimum pada ketinggian 10 m
- winddirection_10m_dominant (°) - Arah angin dominan pada hari tersebut
- city - Nama kota yang tercatat
- class - Jenis/kelompok curah hujan harian - Target Klasifikasi

Data Jam:

- time - Tanggal dan jam pencatatan
- temperature_2m (°C) - Temperatur pada ketinggian 2 m
- relativehumidity_2m (%) - Kelembapan pada ketinggian 2 m

Deskripsi Topik 10

- winddirection_100m (°) – Arah angin pada ketinggian 100 m
- windgusts_10m (km/h) – Kecepatan angin ketika terdapat angin kencang
- etO_fao_evapotranspiration (mm) – Jumlah evapotranspirasi (evaporasi dan transpirasi) pada jam tersebut
- vapor_pressure_deficit (kPa) – Perbedaan tekanan uap air dari udara dengan tekanan uap air ketika udara tersaturasi
- soil_temperature_0_to_7cm (°C) – Rata-rata temperatur tanah pada kedalaman 0-7 cm
- soil_temperature_7_to_28cm (°C) – Rata-rata temperatur tanah pada kedalaman 7-28 cm
- soil_temperature_28_to_100cm (°C) – Rata-rata temperatur tanah pada kedalaman 28-100 cm
- soil_temperature_100_to_255cm (°C) – Rata-rata temperatur tanah pada kedalaman 100-255 cm
- soil_moisture_0_to_7cm (m³/m³) – Rata-rata kelembapan air pada tanah untuk kedalaman 0-7 cm
- soil_moisture_7_to_28cm (m³/m³) – Rata-rata kelembapan air pada tanah untuk kedalaman 7-28 cm

Deskripsi Topik 10

- dewpoint_2m (°C) – Titik embun; suhu ambang udara mengembun
- apparent_temperature (°C) – Temperatur semu yang dirasakan
- pressure_msl (hPa) – Tekanan udara pada ketinggian permukaan air laut rata-rata (mean sea level)
- surface_pressure (hPa) – Tekanan udara pada ketinggian permukaan daerah tersebut
- snowfall (cm) – Jumlah hujan salju pada jam tersebut
- cloudcover_low (%) – Persentase cloud cover pada awan sampai ketinggian 2 km
- cloudcover_mid (%) – Persentase cloud cover pada ketinggian 2-6 km
- cloudcover_high (%) – Persentase cloud cover pada ketinggian di atas 6 km
- windspeed_10m (km/h) – Kecepatan angin pada ketinggian 10 m
- windspeed_100m (km/h) – Kecepatan angin pada ketinggian 100 m
- winddirection_10m (°) – Arah angin pada ketinggian 10 m

Deskripsi Topik 10

- soil_moisture_28_to_100cm (m^3/m^3) – Rata-rata kelembapan air pada tanah untuk kedalaman 28-100 cm
- soil_moisture_100_to_255cm (m^3/m^3) – Rata-rata kelembapan air pada tanah untuk kedalaman 100-255 cm
- city – Nama kota

Gambaran Dataset

1 dtrain.head()

	time	temperature_2m_max (°C)	temperature_2m_min (°C)	apparent_temperature_max (°C)	apparent_temperature_min (°C)	sunrise (iso8601)	sunset (iso8601)	rain_sum (mm)	windspeed_10m_max (km/h)	windgusts_10m_max (km/h)	winddirection_10m_dominant (°)	shortwave_radiation_sum (MJ/m²)	city	class	snowfall	precipitation
0	2018-01-01 00:00:00	33.2	24.3	39.0	28.8	2018-01-01 05:41:14	2018-01-01 18:09:47	1.800000	10.8	23.0	257.02213	23.82	B	Hujan Ringan	0.0	1.800000
1	2018-01-02 00:00:00	33.0	25.1	40.0	29.2	2018-01-02 05:41:45	2018-01-02 18:10:12	2.300000	9.5	19.4	230.53691	22.72	B	Hujan Ringan	0.0	2.300000
2	2018-01-03 00:00:00	31.3	24.9	38.3	30.3	2018-01-03 05:42:14	2018-01-03 18:10:35	11.099999	11.1	23.8	275.63590	22.92	B	Hujan Sedang	0.0	11.099999
3	2018-01-04 00:00:00	31.2	25.9	37.8	31.8	2018-01-04 05:42:43	2018-01-04 18:10:58	16.400000	11.0	31.7	241.11348	17.70	B	Hujan Sedang	0.0	16.400000
4	2018-01-05 00:00:00	29.6	25.4	34.3	29.6	2018-01-05 05:43:12	2018-01-05 18:11:20	10.800000	17.3	33.1	246.59671	9.61	B	Hujan Sedang	0.0	10.800000

1 htrain.head()

	time	temperature_2m (°C)	relativehumidity_2m (%)	dewpoint_2m (°C)	apparent_temperature (°C)	snowfall (cm)	pressure_msl (hPa)	surface_pressure (hPa)	cloudcover_low (%)	cloudcover_mid (%)	...	windgusts_10m (km/h)	soil_temperature_0_to_7cm (°C)	soil_temperature_7_to_28cm (°C)	soil_temperature_28_to_100cm (°C)	soil_temperature_100_to_255cm (°C)	soil_moisture_0_to_7cm (m³/m³)
0	2018-01-01 00:00:00	25.4	91.0	23.9	30.6	0.0	1008.0	1005.8	7.0	10.0	...	9.4	28.1	28.4	28.3	28.7	0.326
1	2018-01-01 01:00:00	25.2	92.0	23.8	30.0	0.0	1007.4	1005.2	16.0	21.0	...	10.1	28.0	28.3	28.3	28.7	0.325
2	2018-01-01 02:00:00	25.1	92.0	23.8	29.8	0.0	1006.9	1004.7	7.0	8.0	...	14.8	28.0	28.2	28.3	28.7	0.324
3	2018-01-01 03:00:00	24.5	93.0	23.3	28.9	0.0	1006.0	1003.8	5.0	4.0	...	15.1	27.4	28.1	28.3	28.7	0.323
4	2018-01-01 04:00:00	24.3	94.0	23.3	28.9	0.0	1005.9	1003.7	2.0	3.0	...	14.0	27.2	28.0	28.3	28.7	0.322

5 rows × 17 columns

Problem Guideline Topik 10

1. Lakukan eksplorasi sederhana pada data dan ceritakan insight yang Anda dapatkan dari data tersebut! (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a) Bagaimana hubungan ke empat kolom mengenai `soil_moisture` dengan jumlah curah hujan harian?
 - b) Apakah ada beberapa kota yang dekat secara geografisnya? Jika ada, pasangan kota mana saja?
 - c) Jika data yang diberikan merupakan data yang nyata, berada di daerah manakah kota K? Jawab permasalahan ini berdasarkan insight yang diperoleh berdasarkan data (Misalnya saja daerah tersebut adalah kukusan, maka dapat dijawab dengan **Indonesia, Jawa Barat, Depok atau Kukusan**. Semakin akurat tingkat daerah yang ditebak maka akan semakin baik)
 - d) Berikan 3 kota yang memiliki rata-rata jumlah curah hujan tahunan tertinggi beserta nilainya!Mahasiswa perlu menambahkan **dua poin eksplorasi sendiri**
2. Buatlah beberapa model yang mungkin akan diperlukan, di antaranya: (dapat ditambah analisis lain yang menurut kelompok Anda akan bermanfaat)
 - a. Terdapat 10 kota dalam data tersebut, buatlah model klustering dengan $k=10$, sedemikian sehingga tiap data akan berada di kota yang sesuai! Tentunya anda tidak diperbolehkan memanfaatkan kolom 'city' dalam pembuatan model ini nantinya
 - b. Buatlah model regresi untuk memprediksi kolom 'rain_sum (mm)' .
 - c. Buatlah model klasifikasi yang dapat memprediksi kelas curah hujan harian atau kolom "class"! Tidak diperbolehkan untuk menggunakan kolom 'rain_sum (mm)' sebagai fitur, namun boleh menghasilkan fitur baru yang mirip dengan 'rain_sum (mm)' dari feature engineering