





Dataset and Modeling Rain in Australia


Kelompok 9 (Marvel)
 [Muhammad Fikri](#),  [Natalia Siringo Ringo](#),  [Satria Triputra W.](#)


 Dataset




 1114

Rain in Australia

Predict next-day rain in Australia

 Joe Young and 1 collaborator • updated 7 months ago (Version 2)

[Data](#) [Tasks \(9\)](#) [Code \(350\)](#) [Discussion \(17\)](#) [Activity](#) [Metadata](#) [Download \(13 MB\)](#) [New Notebook](#) 

 Usability 10.0  License Other (specified in description)  Tags earth and nature, classification, weather and climate, binary classification

Dataset berisikan data observasi cuaca setiap hari selama 10 tahun di seluruh lokasi di Australia.

Dataset Rain in Australia

Data - data yang tersedia dan tipe datanya yaitu :

```
dataset.dtypes
```

```
Date          object
Location       object
MinTemp        float64
MaxTemp        float64
Rainfall       float64
Evaporation    float64
Sunshine       float64
WindGustDir     object
WindGustSpeed   float64
WindDir9am     object
WindDir3pm     object
```

```
WindSpeed9am   float64
WindSpeed3pm   float64
Humidity9am    float64
Humidity3pm    float64
Pressure9am    float64
Pressure3pm    float64
Cloud9am       float64
Cloud3pm       float64
Temp9am        float64
Temp3pm        float64
RainToday      object
RainTomorrow   object
dtype: object
```

Pengertian kolom-kolom di dalam dataset

Nama Kolom	Pengertian	Satuan
Min/Max Temp	Minimum dan maksimum temperatur / suhu udara dalam 24 jam	Derajat Celcius
Rainfall	Tingkat curah hujan untuk menentukan tingkat intensitas hujan dalam waktu 24 jam hingga jam 9 am	milimeter
Evaporation	Intensitas penguapan dalam 24 jam	milimeter
Sunshine	Lamanya penyinaran matahari dalam 24 jam	jam
WindGustDir	Arah hembusan angin terkencang dalam 24 jam	16 arah

Pengertian kolom-kolom di dalam dataset

Nama Kolom	Pengertian	Satuan
WindGustSpeed	Kecepatan hembusan angin terkencang dalam 24 jam. Hembusan angin adalah peningkatan kecepatan angin secara singkat dan tiba-tiba.	Kilometer per jam
WindDir9am/3pm	Arah angin rata-rata lebih dari 10 menit sebelum jam 9 am dan 3 pm	Arah mata angin
WindSpeed9am/3pm	Kecepatan angin yang kita alami rata-rata lebih dari 10 menit sebelum jam 9 am dan 3 pm	Kilometer per jam
Humidity9am/3p	Kelembaban relatif pada jam 9 am dan 3 pm	Persen

Pengertian kolom-kolom di dalam dataset

Nama Kolom	Pengertian	Satuan
Pressure9am/3pm	Tekanan udara pada jam 9 am dan 3 pm. Secara teori, jika tekanan udara rendah, akan ada kemungkinan hujan.	hectopascal
Cloud9am/3pm	Bagian langit yang tertutup awan di jam 9 am dan 3 pm. Berdasarkan WMO (World Meteorological Organization), langit cerah 2 okta, berawan sebagian 2-4 okta, berawan 4-6 okta, penuh awan 6-8 okta, tertutup awan 8 okta.	Perdelapan / okta
Temp9am/3pm	Temperatur pada jam 9 am	Derajat Celcius

Pemahaman dataset

Analisis - analisis yang dapat dilakukan pada Dataset Rain in Australia :

1. Berapa temperatur paling minimum yang terekam?
2. Berapa temperatur paling maksimum yang terekam?
3. Berapa curah hujan terbesar yang terekam?
4. Update data 2017 - 2020
5. Prediksi tanggal kapan akan hujan.
6. Prediksi apakah besok hujan? Berapa persentase kemungkinan hujan besok?
7. Berapa kali terjadi hujan dalam waktu 3 hari berturut di tiap lokasi?
8. Analisis bagaimana fitur wind rainfall temperature di lokasi tertentu dan temukan nilai maksimum dan minimum tiap fitur di tiap lokasi.
9. Menambahkan koordinat geografi lokasi. (Cek ke Tableau)
10. Prediksi persentase tiap fitur (curah hujan, temperatur, kelembaban) di lokasi tertentu.
11. Analisis badai di tiap lokasi. Prediksi kemungkinan badai di tiap lokasi dan arahnya kemana.
12. Analisis peringatan cuaca (*weather alert*) tentang kemungkinan terjadi angin kencang, badai atau tornado. Kemudian, prediksi peringatan cuaca tersebut di tiap lokasi.
13. Dalam 10 tahun terakhir, apakah cuaca ekstrem sering terjadi di tiap lokasi?

Identifikasi Dataset

Pada kasus kali ini kita akan menggunakan dataset dari Kaggle yaitu Rain in Australia, yang di mana pada dataset ini adalah hasil observasi cuaca setiap hari selama 10 tahun yang di mulai sejak tahun 2007 hingga tahun 2017.

Pada dataset Rain in Australia terdapat 49 lokasi, diantaranya:

```
'Albury', 'BadgerysCreek', 'Cobar', 'CoffsHarbour', 'Moree',  
'Newcastle', 'NorahHead', 'NorfolkIsland', 'Penrith', 'Richmond',  
'Sydney', 'SydneyAirport', 'WaggaWagga', 'Williamstown',  
'Wollongong', 'Canberra', 'Tuggeranong', 'MountGinini', 'Ballarat',  
'Bendigo', 'Sale', 'MelbourneAirport', 'Melbourne', 'Mildura',  
'Nhil', 'Portland', 'Watsonia', 'Dartmoor', 'Brisbane', 'Cairns',  
'GoldCoast', 'Townsville', 'Adelaide', 'MountGambier', 'Nuriootpa',  
'Woomera', 'Albany', 'Witchcliffe', 'PearceRAAF', 'PerthAirport',  
'Perth', 'SalmonGums', 'Walpole', 'Hobart', 'Launceston',  
'AliceSprings', 'Darwin', 'Katherine', 'Uluru'], dtype=object)
```


Proses EDA dan Visualisasi

Pada proses **Exploratory Data Analysis (EDA)**, langkah awal yang harus dilakukan adalah melihat **info** tipe data dan menampilkan dimensi dari dataset tersebut.

Pada tahap ini bertujuan untuk mengetahui tipe data dan melihat dimensi dari setiap kolom dari dataset. Dan fungsi **shape** adalah menampilkan dimensi dari dataset, dan dimensi pada dataset Rain in Australia terdapat 145460 baris dan terdapat 23 kolom.

Proses EDA dan Visualisasi

Langkah selanjutnya adalah **checking missing value** pada dataset menggunakan fungsi `isnull().sum()`, yang dimana menampilkan hasil total missing value dari setiap kolom.

Pada tahap ini kita bisa memutuskan kolom mana yang akan di gunakan atau akan dibuang.

Pada case ini kita akan membuang kolom dengan missing value di atas 35%.

	Total	Percentage Null Value %
Sunshine	69835	48.009762
Evaporation	62790	43.166506
Cloud3pm	59358	40.807095
Cloud9am	55888	38.421559
Pressure9am	15065	10.356799
Pressure3pm	15028	10.331363
WindDir9am	10566	7.263853
WindGustDir	10326	7.098859
WindGustSpeed	10263	7.055548
Humidity3pm	4507	3.098446
WindDir3pm	4228	2.906641
Temp3pm	3609	2.481094
RainTomorrow	3267	2.245978
RainToday	3261	2.241853
Rainfall	3261	2.241853
WindSpeed3pm	3062	2.105046
Humidity9am	2654	1.824557
WindSpeed9am	1767	1.214767
Temp9am	1767	1.214767
MinTemp	1485	1.020899
MaxTemp	1261	0.866905
Location	0	0.000000
Date	0	0.000000

Proses EDA dan Visualisasi

Tidak luput juga untuk menampilkan tabel berisikan **statistik descriptive** dari dataset yang kita pakai menggunakan fungsi **Describe**.

Pada tabel di samping kita bisa gunakan hasil statistik descriptive tersebut untuk keperluan missing value nantinya dan masih banyak lagi manfaat dari statistik descriptive.

```
df.describe()
```

	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustSpeed	WindSpeed9am	WindSpeed3pm	Humidity9am	Humidity3pm
count	143975.000000	144199.000000	142199.000000	82670.000000	75625.000000	135197.000000	143693.000000	142398.000000	142806.000000	140953.000000
mean	12.194034	23.221348	2.360918	5.468232	7.611178	40.035230	14.043426	18.662657	68.880831	51.539111
std	6.398495	7.119049	8.478060	4.193704	3.785483	13.607062	8.915375	8.809800	19.029164	20.795900
min	-8.500000	-4.800000	0.000000	0.000000	0.000000	6.000000	0.000000	0.000000	0.000000	0.000000
25%	7.600000	17.900000	0.000000	2.600000	4.800000	31.000000	7.000000	13.000000	57.000000	37.000000
50%	12.000000	22.600000	0.000000	4.800000	8.400000	39.000000	13.000000	19.000000	70.000000	52.000000
75%	16.900000	28.200000	0.800000	7.400000	10.600000	48.000000	19.000000	24.000000	83.000000	66.000000
max	33.900000	48.100000	371.000000	145.000000	14.500000	135.000000	130.000000	87.000000	100.000000	100.000000

```
df.describe(include=[object])
```

	Date	Location	WindGustDir	WindDir9am	WindDir3pm	RainToday	RainTomorrow
count	145460	145460	135134	134894	141232	142199	142193
unique	3436	49	16	16	16	2	2
top	2014-05-07	Canberra	W	N	SE	No	No
freq	49	3436	9915	11758	10838	110319	110316

Proses EDA dan Visualisasi

Pada tahap visualisasi ini kita akan menampilkan Distribusi dan korelasi dari setiap kolom.

Korelasi adalah proses mengukur hubungan antara dua set nilai.

Tabel di samping adalah 5 list dengan korelasi tertinggi, dan kita akan membuang kolom Temp9am dan Temp3am.

Kolom 1	Kolom 2	Korelasi
Temp9am	MinTemp	0.90
	MaxTemp	0.89
Temp3pm	MinTemp	0.71
	MaxTemp	0.98
Pressure9am	Pressure3pm	0.96

Proses EDA dan Visualisasi

Pada visualisasi di bawah kita akan menghapus kolom Temp9am dan Temp3pm

Before

	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustSpeed	WindSpeed9am	WindSpeed3pm	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm	Temp9am	Temp3pm
MinTemp	1	0.74	0.1	0.47	0.073	0.18	0.18	0.18	-0.23	0.0061	-0.45	-0.46	0.079	0.022	0.9	0.71
MaxTemp	0.74	1	-0.075	0.59	0.47	0.068	0.014	0.05	-0.5	-0.51	-0.33	-0.43	-0.29	-0.28	0.89	0.98
Rainfall	0.1	-0.075	1	-0.064	-0.23	0.13	0.087	0.058	0.22	0.26	-0.17	-0.13	0.2	0.17	0.011	-0.08
Evaporation	0.47	0.59	-0.064	1	0.37	0.2	0.19	0.13	-0.5	-0.39	-0.27	-0.29	-0.18	-0.18	0.55	0.57
Sunshine	0.073	0.47	-0.23	0.37	1	-0.035	0.0055	0.054	-0.49	-0.63	0.042	-0.02	-0.68	-0.7	0.29	0.49
WindGustSpeed	0.18	0.068	0.13	0.2	-0.035	1	0.61	0.69	-0.22	-0.026	-0.46	-0.41	0.072	0.11	0.15	0.033
WindSpeed9am	0.18	0.014	0.087	0.19	0.0055	0.61	1	0.52	-0.27	-0.032	-0.23	-0.18	0.025	0.055	0.13	0.0046
WindSpeed3pm	0.18	0.05	0.058	0.13	0.054	0.69	0.52	1	-0.15	0.016	-0.3	-0.26	0.053	0.025	0.16	0.028
Humidity9am	-0.23	-0.5	0.22	-0.5	-0.49	-0.22	-0.27	-0.15	1	0.67	0.14	0.19	0.45	0.36	-0.47	-0.5
Humidity3pm	0.0061	-0.51	0.26	-0.39	-0.63	-0.026	-0.032	0.016	0.67	1	-0.028	0.052	0.52	0.52	-0.22	-0.56
Pressure9am	-0.45	-0.33	-0.17	-0.27	0.042	-0.46	-0.23	-0.3	0.14	-0.028	1	0.96	-0.13	-0.15	-0.42	-0.29
Pressure3pm	-0.46	-0.43	-0.13	-0.29	-0.02	-0.41	-0.18	-0.26	0.19	0.052	0.96	1	-0.061	-0.085	-0.47	-0.39
Cloud9am	0.079	-0.29	0.2	-0.18	-0.68	0.072	0.025	0.053	0.45	0.52	-0.13	-0.061	1	0.6	-0.14	-0.3
Cloud3pm	0.022	-0.28	0.17	-0.18	-0.7	0.11	0.055	0.025	0.36	0.52	-0.15	-0.085	0.6	1	-0.13	-0.32
Temp9am	0.9	0.89	0.011	0.55	0.29	0.15	0.13	0.16	-0.47	-0.22	-0.42	-0.47	-0.14	-0.13	1	0.86
Temp3pm	0.71	0.98	-0.08	0.57	0.49	0.033	0.0046	0.028	-0.5	-0.56	-0.29	-0.39	-0.3	-0.32	0.86	1

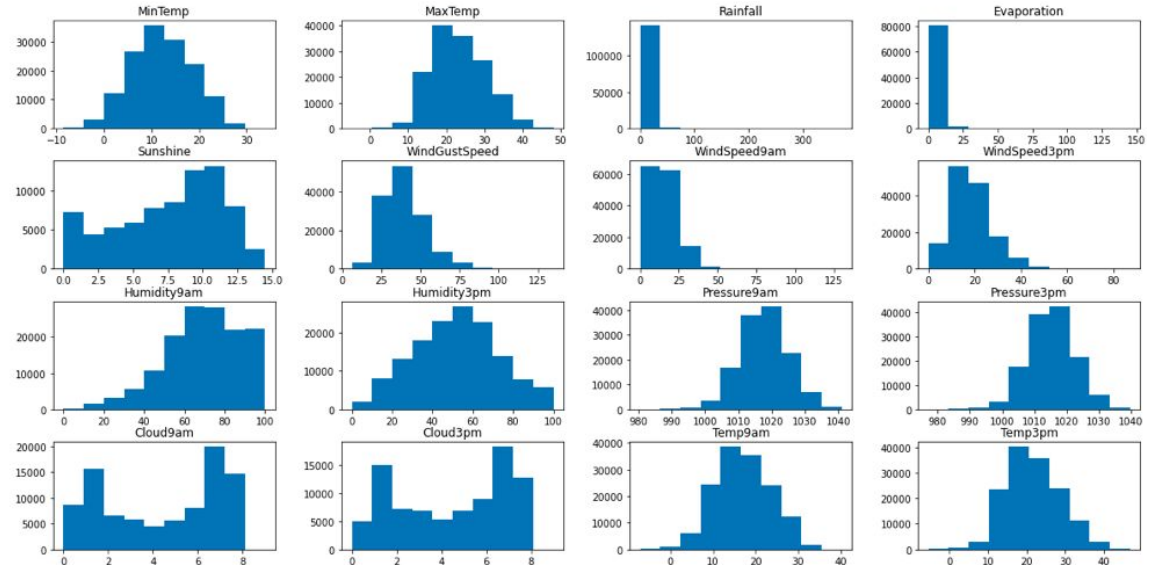
After

	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustSpeed	WindSpeed9am	WindSpeed3pm	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Cloud9am	Cloud3pm	Temp9am	Temp3pm
MinTemp	1	0.74	0.1	0.47	0.073	0.18	0.18	0.18	-0.23	0.0061	-0.45	-0.46	0.079	0.022	0.9	0.71
MaxTemp	0.74	1	-0.075	0.59	0.47	0.068	0.014	0.05	-0.5	-0.51	-0.33	-0.43	-0.29	-0.28	0.89	0.98
Rainfall	0.1	-0.075	1	-0.064	-0.23	0.13	0.087	0.058	0.22	0.26	-0.17	-0.13	0.2	0.17	0.011	-0.08
Evaporation	0.47	0.59	-0.064	1	0.37	0.2	0.19	0.13	-0.5	-0.39	-0.27	-0.29	-0.18	-0.18	0.55	0.57
Sunshine	0.073	0.47	-0.23	0.37	1	-0.035	0.0055	0.054	-0.49	-0.63	0.042	-0.02	-0.68	-0.7	0.29	0.49
WindGustSpeed	0.18	0.068	0.13	0.2	-0.035	1	0.61	0.69	-0.22	-0.026	-0.46	-0.41	0.072	0.11	0.15	0.033
WindSpeed9am	0.18	0.014	0.087	0.19	0.0055	0.61	1	0.52	-0.27	-0.032	-0.23	-0.18	0.025	0.055	0.13	0.0046
WindSpeed3pm	0.18	0.05	0.058	0.13	0.054	0.69	0.52	1	-0.15	0.016	-0.3	-0.26	0.053	0.025	0.16	0.028
Humidity9am	-0.23	-0.5	0.22	-0.5	-0.49	-0.22	-0.27	-0.15	1	0.67	0.14	0.19	0.45	0.36	-0.47	-0.5
Humidity3pm	0.0061	-0.51	0.26	-0.39	-0.63	-0.026	-0.032	0.016	0.67	1	-0.028	0.052	0.52	0.52	-0.22	-0.56
Pressure9am	-0.45	-0.33	-0.17	-0.27	0.042	-0.46	-0.23	-0.3	0.14	-0.028	1	0.96	-0.13	-0.15	-0.42	-0.29
Pressure3pm	-0.46	-0.43	-0.13	-0.29	-0.02	-0.41	-0.18	-0.26	0.19	0.052	0.96	1	-0.061	-0.085	-0.47	-0.39
Cloud9am	0.079	-0.29	0.2	-0.18	-0.68	0.072	0.025	0.053	0.45	0.52	-0.13	-0.061	1	0.6	-0.14	-0.3
Cloud3pm	0.022	-0.28	0.17	-0.18	-0.7	0.11	0.055	0.025	0.36	0.52	-0.15	-0.085	0.6	1	-0.13	-0.32
Temp9am	0.9	0.89	0.011	0.55	0.29	0.15	0.13	0.16	-0.47	-0.22	-0.42	-0.47	-0.14	-0.13	1	0.86
Temp3pm	0.71	0.98	-0.08	0.57	0.49	0.033	0.0046	0.028	-0.5	-0.56	-0.29	-0.39	-0.3	-0.32	0.86	1

Proses EDA dan Visualisasi

Gambar di samping adalah hasil visualisasi Distribusi.

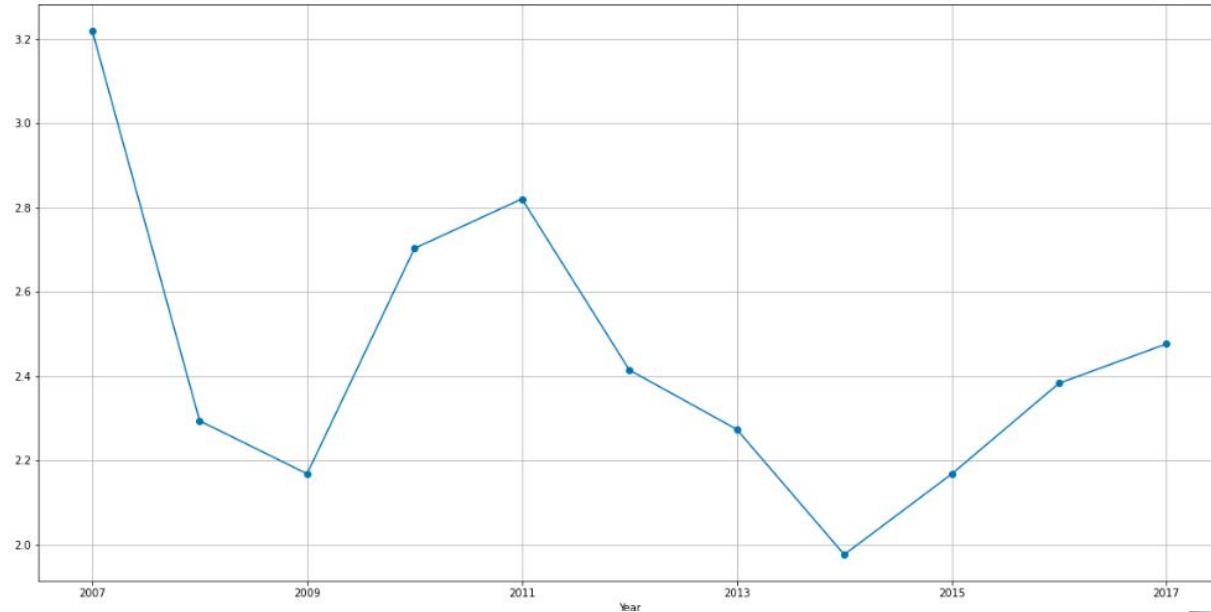
Distribusi data adalah suatu fungsi yang menunjukkan semua nilai dari sebuah data dan seberapa sering nilai tersebut terjadi.



Proses EDA dan Visualisasi

Visualisasi di samping adalah rata-rata curah hujan selama 10 tahun yang dimulai dari tahun 2007 hingga 2017

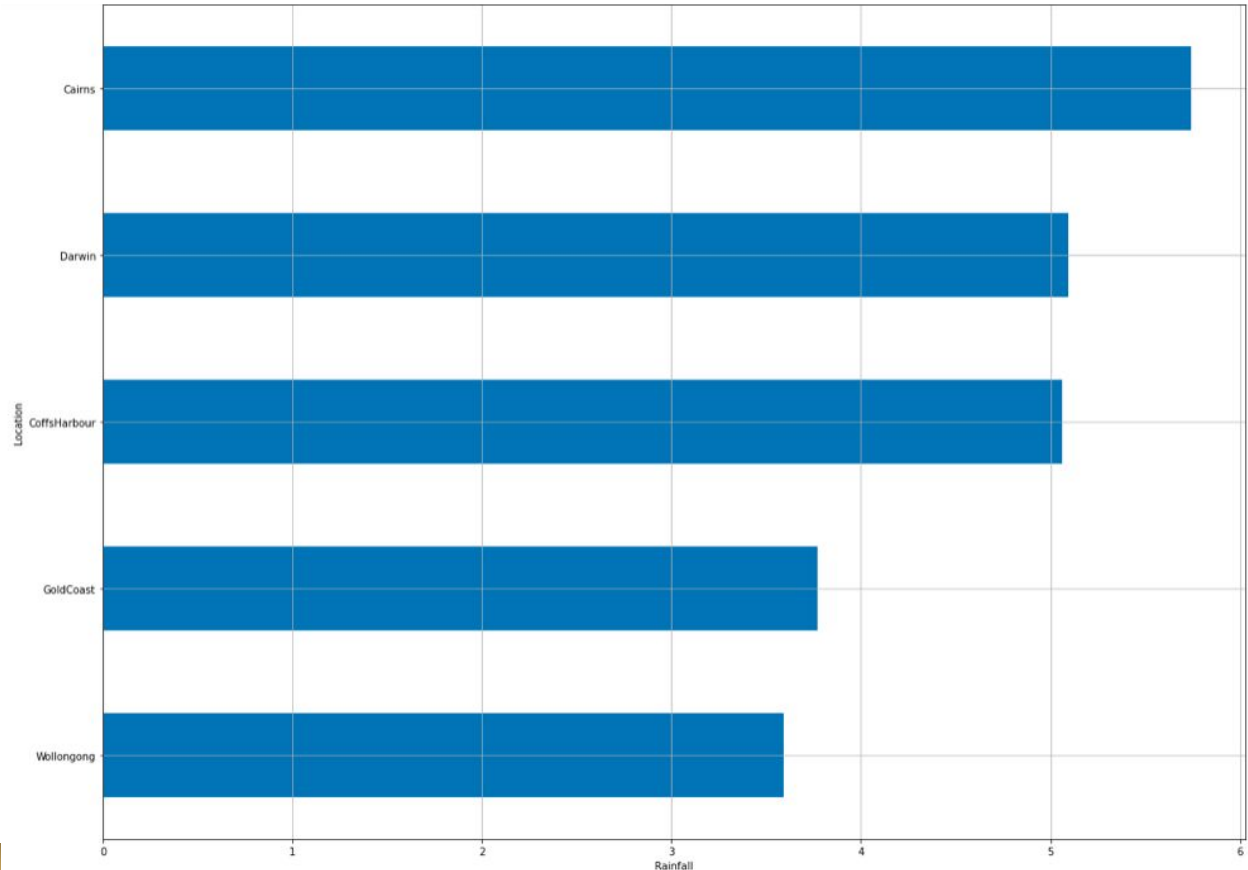
Kita bisa lihat pada visualisasi di samping bahwa pada tahun 2008, 2009 dan 2014 curah hujan di Australia mengalami penurunan yang cukup signifikan yang dimana disebabkan oleh pergantian musim panas di bulan desember - februari.



Proses EDA dan Visualisasi

Setelah kita mengetahui nama-nama lokasi, langkah selanjutnya kita bisa mevisualisasikan curah hujan terbesar berada di wilayah mana.

Dan pada visualisasi di samping terdapat 5 wilayah dengan rata-rata curah hujan terbesar di atas satuan milimeter 3, diantaranya adalah **Cairns**, **Darwin**, **CoffsHarbour**, **GoldCoast**, dan **Wollongong**



Proses Preprocessing

Pada proses **PreProcessing** hal utama yang harus dilakukan adalah **Handling missing value**.

Pada tahap ini kita akan mengisi baris yang kosong / null dengan **mean()** dan **mode()**.

Before

```
MinTemp      0
MaxTemp      0
Rainfall     0
WindGustSpeed 6079
WindSpeed9am  0
WindSpeed3pm  0
Humidity9am   0
Humidity3pm   0
Pressure9am   12119
Pressure3pm   12119
Temp9am       0
Temp3pm       0
dtype: int64
```

After

```
MinTemp      0
MaxTemp      0
Rainfall     0
WindGustSpeed 0
WindSpeed9am  0
WindSpeed3pm  0
Humidity9am   0
Humidity3pm   0
Pressure9am   0
Pressure3pm   0
Temp9am       0
Temp3pm       0
dtype: int64
```

Proses Preprocessing

Selain handling missing value, yang harus di perhatikan adalah merubah atau mereplace tipe data Categorical menjadi Numerik pada kasus ini kita menggunakan fungsi **value_counts()**, yang dimana tujuan nya adalah agar Machine Learning bisa memproses data tersebut.

	MinTemp	MaxTemp	Rainfall	WindGustSpeed	WindSpeed9am	WindSpeed3pm	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	Temp9am	Ten
0	13.4	22.900000	0.6	44.000000	20.0	24.0	71.0	22.0	1007.7	1007.1	16.9	
1	7.4	25.100000	0.0	44.000000	4.0	22.0	44.0	25.0	1010.6	1007.8	17.2	
2	12.9	25.700000	0.0	46.000000	19.0	26.0	38.0	30.0	1007.6	1008.7	21.0	
3	9.2	28.000000	0.0	24.000000	11.0	9.0	45.0	16.0	1017.6	1012.8	18.1	
4	17.5	32.300000	1.0	41.000000	7.0	20.0	82.0	33.0	1010.8	1006.0	17.8	
...
145455	2.8	23.400000	0.0	31.000000	13.0	11.0	51.0	24.0	1024.6	1020.3	10.1	
145456	3.6	25.300000	0.0	22.000000	13.0	9.0	56.0	21.0	1023.5	1019.1	10.9	
145457	5.4	26.900000	0.0	37.000000	9.0	9.0	53.0	24.0	1021.0	1016.8	12.5	
145458	7.8	27.000000	0.0	28.000000	13.0	7.0	51.0	24.0	1019.4	1016.5	15.1	
145459	14.9	30.383195	0.0	41.268373	17.0	17.0	62.0	36.0	1020.2	1017.9	15.0	

145460 rows x 16 columns

Proses Preprocessing

Setelah semua kolom menjadi numerik, selanjutnya adalah proses **Standarisasi**. Secara teknis, standarisasi memusatkan dan menormalkan data dengan mengurangi mean dan membaginya dengan standar deviasi. Pada tahap ini saya menggunakan fungsi **StandardScaler** dari package **sklearn**

MinTemp	MaxTemp	Rainfall	WindGustSpeed	WindSpeed9am	WindSpeed3pm	Humidity9am	Humidity3pm	Pressure9am	Pressure3pm	WindGustDir	WindDir
0.189479	-0.043902	-0.210380	0.299901	0.673183	0.616456	0.108840	-1.442551	-1.476793	-1.222913	2.434305	-0.308
-0.751913	0.266079	-0.281947	0.299901	-1.129806	0.387759	-1.320713	-1.296626	-1.045963	-1.117903	-0.443402	-0.192
0.111029	0.350619	-0.281947	0.452211	0.560496	0.845154	-1.638391	-1.053417	-1.491649	-0.982890	-0.247298	-0.308
-0.469496	0.674691	-0.281947	-1.223205	-0.340998	-1.098776	-1.267767	-1.734402	-0.006030	-0.367833	-0.711994	0.646
0.832763	1.280564	-0.162669	0.071435	-0.791746	0.159061	0.691250	-0.907491	-1.016251	-1.387928	2.434305	-0.828
...
-1.473647	0.026548	-0.281947	-0.690118	-0.115625	-0.870078	-0.950088	-1.345268	1.033904	0.757273	-0.220415	0.646
-1.348128	0.294259	-0.281947	-1.375515	-0.115625	-1.098776	-0.685356	-1.491193	0.870485	0.577256	-0.835129	0.646
-1.065711	0.519700	-0.281947	-0.233186	-0.566372	-1.098776	-0.844195	-1.345268	0.499081	0.232223	-0.188731	0.646
-0.689154	0.533790	-0.281947	-0.918584	-0.115625	-1.327473	-0.950088	-1.345268	0.261382	0.187219	-0.163528	0.102
0.424827	1.010485	-0.281947	0.091873	0.335123	-0.183985	-0.367678	-0.761566	0.380231	0.397239	2.434305	-0.965

rows x 15 columns

Proses Preprocessing

Langkah terakhir dalam proses PreProcessing adalah **Splitting Dataset**. Yang dimana langkah ini bertujuan untuk membagi data menjadi data training dan data testing.

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(df_x, df_cat['RainTomorrow'], test_size=0.2, random_state=50)
```

Proses Preprocessing

Jika dataset yang kita miliki terdapat data yang tidak seimbang, maka langkah yang harus kita ambil adalah menyetarakannya dengan menggunakan teknik **resampling** dari package **imblearn** dan menggunakan fungsi **SMOTE**.

```
# Oversampled

from imblearn.over_sampling import SMOTE
from collections import Counter

sm = SMOTE(random_state=37)
X_train_res, y_train_res = sm.fit_resample(X_train, y_train)

print('Sebelum {}'.format(Counter(y_train)))
print('Sesudah {}'.format(Counter(y_train_res)))
```

```
Sebelum Counter({0: 90784, 1: 25584})
Sesudah Counter({1: 90784, 0: 90784})
```

Pada gambar di samping, saat sebelum di Resampling kolom No(0) memiliki 90784 baris dan kolom Yes(1) 25584 baris.

Setelah kita Resampling menggunakan fungsi SMOTE maka kolom Yes(1) dan kolom No(0) menjadi seimbang.

Proses Modelling

Pada tahap **modelling** kita akan menggunakan beberapa algoritma Machine Learning, diantaranya adalah **Decision Tree, Logistic Regression, KNN, dan SVM**. Tujuannya adalah untuk membandingkan Algoritma Klasifikasi manakah yang menunjukkan akurasi yang tinggi.

Pada algoritma di samping kita bisa lihat bahwa Algoritma **Decision Tree** memiliki akurasi yang cukup tinggi.

Algoritma Machine Learning	Percentage Accuracy %
Decision Tree	80.68 %
SVM (Support Vector Machine)	78.53 %
Logistic Regression	78.31 %
KNN (K-Nearest Neighbors)	77.47 %

Special Thanks to :



Slide template by SlideCarnival