In [ ]:
```
#Importation et affichage
```

In [35]:
```
pwd
```

Out[35]:
```
'/Users/Yaha/Desktop/CAPSTONE 2'
```

In [36]:
```
import pandas as pd
```

In [37]:
```
df_weather = pd.read_csv('/Users/Yaha/Desktop/CAPSTONE 2/Dataset/Weather/weather.csv', sep = ';')
```

In [38]:
```
df_weather
```

Out[38]:

	Day	time_of_day	weather	severity	degrees
0	05/23/22	Morning	Rain	Heavy	14
1	05/23/22	Afternoon	Rain	Light	18
2	05/23/22	Evening	Clear	None	16
3	05/24/22	Morning	Rain	Heavy	11
4	05/24/22	Afternoon	Rain	Heavy	12
...	...	...	...	...	...
79	06/18/22	Afternoon	Clear	None	26
80	06/18/22	Evening	Clear	None	24
81	06/19/22	Morning	Clear	None	24
82	06/19/22	Afternoon	Clear	None	28
83	06/19/22	Evening	Clear	None	26

84 rows × 5 columns

In [ ]:
```
#Informations sur la dataframe
```

In [46]:
```
df_weather.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 84 entries, 0 to 83
Data columns (total 5 columns):
 #   Column       Non-Null Count  Dtype 
---  ------       --------------  ----- 
 0   tour_date    84 non-null     datetime64[ns]
 1   time_of_day  84 non-null     object
 2   weather      84 non-null     object
 3   severity     84 non-null     object
 4   degrees      84 non-null     int64 
dtypes: datetime64[ns](1), int64(1), object(3)
memory usage: 3.4+ KB
```

In [48]:
```
#Détection des valeurs manquantes
df_weather.isnull().sum()
```

Out[48]:
```
tour_date      0
time_of_day    0
weather        0
severity       0
degrees        0
dtype: int64
```

In [49]:
```
#Nombre de doublons
print(df_weather.duplicated().sum())
```

```
0
```

In [ ]:
```
#Changement de la colonne 'Day' renommé en 'tour_date'
```

In [39]:
```
df_weather = df_weather.rename(columns={'Day':'tour_date'})
```

In [40]:
```
df_weather
```

Out[40]:

	tour_date	time_of_day	weather	severity	degrees
0	05/23/22	Morning	Rain	Heavy	14
1	05/23/22	Afternoon	Rain	Light	18
2	05/23/22	Evening	Clear	None	16
3	05/24/22	Morning	Rain	Heavy	11
4	05/24/22	Afternoon	Rain	Heavy	12
...	...	...	...	...	...
79	06/18/22	Afternoon	Clear	None	26
80	06/18/22	Evening	Clear	None	24
81	06/19/22	Morning	Clear	None	24
82	06/19/22	Afternoon	Clear	None	28
83	06/19/22	Evening	Clear	None	26

84 rows × 5 columns

In [ ]:
```
#Uniformisation du format 'tour_date'
```

In [41]:
```
df_weather['tour_date'] = pd.to_datetime(df_weather['tour_date'])
```

In [42]:
```
df_weather
```

Out[42]:

	tour_date	time_of_day	weather	severity	degrees
0	2022-05-23	Morning	Rain	Heavy	14
1	2022-05-23	Afternoon	Rain	Light	18
2	2022-05-23	Evening	Clear	None	16
3	2022-05-24	Morning	Rain	Heavy	11
4	2022-05-24	Afternoon	Rain	Heavy	12
...	...	...	...	...	...
79	2022-06-18	Afternoon	Clear	None	26
80	2022-06-18	Evening	Clear	None	24
81	2022-06-19	Morning	Clear	None	24
82	2022-06-19	Afternoon	Clear	None	28
83	2022-06-19	Evening	Clear	None	26

84 rows × 5 columns

In [45]:
```
df_weather.to_csv('Output\\weather.csv',index=False)
```