UOA MSDS C2 14/05/2022

Uber Pre-assignment

(3) `sum( df.duplicated(subset="Req_id")) == 0`

(5) `df.isnull().sum()`

(7) `df['R_id'].isnull().values.any()` ——— `df['R_id'].isnull().sum()`

(14) `req_hr = df['R_TS'].dt.hour`

apply value_counts()

`df['req_hr'] = req_hr.`

(16) `sns.facplot(x='reqhr', hue='Status', data=df, kind=count)`

(1a)

```
def time_per(x):
    if x < 5
        return "pre Morning"
    elif 5 ≤ x < 10:
        Morning rush
    elif 10 ≤ x ≤ 17:
        Day time
    elif 17 ≤ x < 22:
        Evening Rush
    else:
        Late evening
```

```
df['time slot'] = df.req_hr .apply( lambda.. )
```

23

```
df_morn_rush = df[df.timeslot == 'Morning rsh']
sns.countplot(x='PUP', hue=Status, data=df_mrg rsh)
```

24 `df_city_cancelled = df_morn_rush.loc[`

`(df_morn_rush[PUP] == Airport) &`

`(df_morn_rush[Status] == Cancelled)]`

`len(df_city_cancelled.index)`

30 `df_even_rush = df[df.timeslot == even rush]`

`sns.countplot(x=PUP, hue=Status, data= )`