นับ rows,cols ใช้ df.shape[0], df.shape[1]

นับ NaN -> ใช้ df.yourCols.isna().sum()

นับไม่ NaN ใช้ df.yourCols.notna().sum()

iloc ใช้ได้เฉพาะ integer loc, loc ใช้ปนกันได้ โดยให้ loc['เงื่อนไขของ rows' : 'เงื่อนไขของ cols']

เราต้องใส่เป็นชนิด value เช่น df['xxxx'] หรือ df.xxxxx

เราต้องใช้ & และ | ในการและ, หรือ เท่านั้น

df.loc[: , :]

df.loc[(df['a'] > df['b']) & (df['c'] < df['d']), ['a', 'b', 'c']]

df.loc[df['a'] == df.a.max() , :] <- loc แถวที่ค่าจาก col a มากที่สุด

นับจำนวนจาก cols ที่สนใจ df.yourCols.str.contains('xxxxxx').sum()

Loc จาก cols ที่สนใจโดยมี NaN ด้วยให้ใช้ loc คู่กับ df.yourCols.str.contains('xxxxxx').fillna(False) <- ไม่มอง NaN

เช่น dfN = air\_crash.loc[air\_crash.Location.str.contains('Thailand').fillna(False), : ]

จะ sort ค่า ใส่เป็น string ตรงๆได้

ให้ df.sort\_values([] , ascending=[])

เช่น air\_crash.sort\_values(['Fatalities Percent','Aboard'], ascending=[False, False])

Seaborn

displot

sns.displot(df.yourCols, kind='??', height=0)

e.g. sns.displot(usa\_player.shots, kind='hist',height=3)

e.g. sns.displot(players.passes, kind='kde', height=3)

A graph with blue bars

Description automatically generatedkind = ‘hist’ -> Histrogram

kind = ‘kde’ -> Probability density Function

A blue line graph with numbers

Description automatically generated

displot among multiple datas

sns.displot(filtered\_data, x=’colname1’, kine=’kde’, height=0, hue=’colname2’)

e.g. sns.displot(data, x='passes', kind='kde', height=3, hue='team')

A diagram of a line graph

Description automatically generated

Relplot

Assume filtered\_data is your filtered interested data (e.g. only usa team)

Relplot defaults as ‘scatter’

sns.relplot(filtered\_data, x='x1', y='y1', hue='colname1', style='colname2')

e.g. sns.relplot(players, x='passes', y='shots', height=3, aspect=1.5)

A graph of blue dots

Description automatically generated

data\_aggregated = data2.groupby('team').agg({'passes': 'sum', 'shots': 'sum'})

sns.relplot(data\_aggregated, x='passes', y='shots', hue='team', kind='scatter') <- kind defaults as scatter

A graph with colored dots

Description automatically generated

Multiple relplot

sns.relplot(data, x=’x1’, y=’y1’, row=’colname1’, col=’colname2’)

e.g. sns.relplot(data, x='tackles', y='shots', row='position', col='team')

A screenshot of a graph

Description automatically generated

Catplot

sns.catplot(filtered\_data, x='x1', y='y1', height=0, kind='??', hue='colname1')

A group of graphs on a graph paper

Description automatically generatedA graph of different colored squares

Description automatically generatede.g. sns.catplot(data, x='team', y='passes', height=3, kind='box', hue='team')