Apuntes Modulo 3

178679

Para comparar dos columnas de datos podemos usar

Scatter(x,y)

Se acostusmbra poner en el eje de las x, la variable base

Y en el eje de las y la variable que queremos predecir

Para agrupar datos

Df.Groupby([‘lista de columnas en string’], as index=false).mean() (para las columnas no agrupadas)

Pivot

.pivot (index =’primer característica row, columns=’columna para pivote’)

.plt.pcolor(df,cmap)

Plt.colorbar()

.plt.show()

Anova

F-test, radio de variación (menor, diferencia, menor correlación)

p-value\_ confianza

.groupb([‘columna’])

Correalciones

Interdependencia entre variables

Sns.regplot(x, y, df)

Plt.ylim(0,)

Estadísticas de correlaciones

Coeficiente, cerca de 1 fuere positiva, -1 fuerte negativa, 0 no correlacion

P-value

<0.001 fuerza

< 0.05 moderada

<0.1 debil

>0.1 Sin correlacion