

### في الطلب الاول :

حددنا عدد الصفوف والاعمدة عن طريق df.shape

وحددنا نوع البيانات في كل عمود df.dtypes

وعرفنا عدد البيانات المفقودة في كل عمود df.isnull().sum()

### في الطلب الثاني:

قمنا بحساب mean, median, mode, minimum, and maximum salary عن طريق توابع جاهزة

وحسبنا مجال الرواتب من الفرق بين اكبر واصغر راتب و كذلك حسبنا الانحراف المعياري عن طريق

df['TotalPay'].std()

### في الطلب الثالث :

ملئ القيم الفارغة :

حذفنا Notes , Status لان كل قيمهم فارغة

عمود Benefits القيم الفارغة ملئناها بقيم فارغة

عمود BasePay القيم الفارغة قمنا بحسابها عن طريق الفرق بين الاجر الإجمالي و الفوائد

### في الطلب الرابع :

1. قمنا برسم بياني يوضح توزيع الرواتب بناءً على عمود "TotalPay". يمثل المحور x قيم الرواتب، ويمثل المحور

y تكرار أو عدد الرواتب

2. المطلوب توزيع الموظفين في الاقسام: قمنا بحساب مخطط دائري يصور نسبة الموظفين في كل قسم بناءً على

عمود "JobTitle" في مجموعة البيانات يتم تمثيل كل قسم بشريحة دائرية، ويتم عرض النسبة المئوية للموظفين

في كل قسم على الرسم البياني.

### في الطلب الخامس :

تقوم groupby(['JobTitle', 'Year']) بتجميع البيانات حسب العمودين 'JobTitle' و 'Year'. يقوم الجزء

agg({'TotalPay': 'mean', 'BasePay': 'mean'}) بحساب متوسط العمودين 'TotalPay' و 'BasePay' لكل مجموعة.

تعمل الدالة Pivot\_table(index='JobTitle', columns='Year', values='TotalPay') على إعادة ترتيب البيانات

لمقارنة متوسط الرواتب عبر مجموعات مختلفة. يصبح عمود "JobTitle" هو الفهرس، ويتحول عمود "السنة" إلى الأعمدة،

ويصبح عمود "TotalPay" هو القيم. تعرض طباعة (average\_salary\_comparison) الجدول الناتج الذي يقارن متوسط

الرواتب عبر المسميات الوظيفية والسنوات المختلفة. في النهاية سنحصل على جدول يوضح متوسط الرواتب للمسميات

الوظيفية المختلفة في سنوات مختلفة، مما يسمح بمقارنة متوسط الرواتب عبر المجموعات المختلفة.

### في الطلب السادس :

سوف تحصل على مخطط مبعثر يصور العلاقة بين أعمدة الراتب والسنة، وسيتم عرض قيمة الارتباط في عنوان المخطط

اخترت عمود السنة لمعرفة ارتباطه بالراتب ليساعد بمعرفة الاسعار خلال كل سنة وتغيرها في السنوات ومعرفة مدى ارتباط السنة بالراتب.