

```

import pandas as pd

pd.set_option("display.max.rows",None)

pd.set_option("max_colwidth",None)

#Q1

print ("Q1")

df = pd.read_csv (r'ds_salaries.csv')

print (df) ##print out csv file

print("#####")

##Question2

print ("Q2")

df = pd.read_csv (r'ds_salaries.csv')

df.groupby("work_year").mean()["salary_in_usd"] ##group teh data by year and calcualte the avaerage salary in USD

print(df.groupby("work_year").mean()["salary_in_usd"])

print("#####")

#Question3

print ("Q3")

min_year = df.groupby("work_year").mean()["salary_in_usd"].head(1)##group the data by year and calcualte the avaerage salary in USD
and find the lowest one

max_year = df.groupby("work_year").mean()["salary_in_usd"].tail(1)##group the data by year and calcualte the avaerage salary in USD
and find the highest one

print("The highest average salary year ")

print(max_year)

print()

print("The lowest average salary year ")

print(min_year)

print("#####")

#Question4

print ("Q4")

Avg_lv_jt = df.groupby(['experience_level','job_title']).mean()["salary_in_usd"]##group the data by year and experience level then
calculate the average salary in USD

print(Avg_lv_jt)

print("#####")

#Quesion5

print ("Q5")

job_title_max = df.groupby("job_title").max()["salary_in_usd"]##group teh data by year and calcualte the avaerage salary in USD and
find each tile highest salary in USD

job_title_min = df.groupby("job_title").min()["salary_in_usd"]##group teh data by year and calcualte the avaerage salary in USD and find
each tile lowest salary in USD

print("Maxium:")

print(job_title_max)

print("Minium:")

print(job_title_min)

print("#####")

#Question6

```

```

print("Q6")
Avg_av_jt = df.groupby(['work_year','job_title']).mean()["salary_in_usd"]##group teh data by year and job title then calualte the
avaerage salary in USD
print(Avg_av_jt)
print("#####")
#Q7
print("Q7")
df.head()
a = df.groupby(['job_title','work_year'],as_index=False)['salary_in_usd'].mean()##group teh data by year and job title then calualte
the avaerage salary in USD
s = pd.DataFrame(a) ##set a new data frame
b = s.groupby(['job_title'],as_index=False)['salary_in_usd'].max()##find max
c = s.groupby(['job_title'],as_index=False)['salary_in_usd'].min()##findd min
df1 = pd.DataFrame(b)
df2 = pd.DataFrame(c)
df3 = df2.copy()
df3['difference'] = df1['salary_in_usd'] - df3['salary_in_usd']##function for calculate the change
df3 = df3[df3['difference'] > 0] ##becuse some job title just apeear once, so we need set the non zero difference
n1 = df3['job_title'][df3['difference']==df3['difference'].max()]
n2 = df3['job_title'][df3['difference']==df3['difference'].min()]
print("Highest chang job title")
print(n1)
print()
print("Lowest chang job title")
print(n2)
print("#####")
#Question8
print("Q8")
average_ratio = df.groupby("remote_ratio").mean()["salary_in_usd"]##group the data remote ratio and then calualte the avaerage
salary in USD
print(average_ratio)
print("There is/are:",len(average_ratio),"type/s of ratio")
print("#####")
#Question9
print("Q9")
upper_bound_salary = df.groupby("company_location").max()["salary_in_usd"].max()##group the data location then find highesrt salary
lower_bound_salary = df.groupby("company_location").min()["salary_in_usd"].min()##group the data location then find lowest salary
print("Highest Loaction:")
print(df.loc[df['salary_in_usd'] == upper_bound_salary, 'company_location'])##return the highest salary location
print()
print("Lowest Location:")
print(df.loc[df['salary_in_usd'] == lower_bound_salary, 'company_location'])##return the lowest salary location
print("#####")
#Question10

```

```
print("Q10")
```

```
print("I would not like to change my resume right now")
```