In [672…

```
import pandas as pd
import numpy as np
```

Question #1. load the dataset csv file as a dataframe using Pandas

In [673…

```
df = pd.read_csv("ds_salaries.csv")
print(df.shape)
print(df.info())
# print(df.head)
# print(df.tail)
print(df.isnull())
print(df.dtypes)
df.columns
df.index
```

```
(607, 12)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 607 entries, 0 to 606
Data columns (total 12 columns):
 #   Column              Non-Null Count  Dtype 
---  ------              --------------  ----- 
 0   Unnamed: 0          607 non-null    int64 
 1   work_year           607 non-null    int64 
 2   experience_level    607 non-null    object
 3   employment_type     607 non-null    object
 4   job_title           607 non-null    object
 5   salary              607 non-null    int64 
 6   salary_currency     607 non-null    object
 7   salary_in_usd       607 non-null    int64 
 8   employee_residence  607 non-null    object
 9   remote_ratio        607 non-null    int64 
 10  company_location    607 non-null    object
 11  company_size        607 non-null    object
dtypes: int64(5), object(7)
memory usage: 57.0+ KB
None
     Unnamed: 0  work_year  experience_level  employment_type  job_title  \
0         False      False             False            False      False   
1         False      False             False            False      False   
2         False      False             False            False      False   
3         False      False             False            False      False   
4         False      False             False            False      False   
..          ...        ...               ...              ...        ...   
602       False      False             False            False      False   
603       False      False             False            False      False   
604       False      False             False            False      False   
605       False      False             False            False      False   
606       False      False             False            False      False   

     salary  salary_currency  salary_in_usd  employee_residence  remote_ratio  \
0     False            False          False               False         False   
1     False            False          False               False         False   
2     False            False          False               False         False   
```

```
3     False            False          False              False         False
4     False            False          False              False         False
..      ...              ...            ...                ...           ...
602   False            False          False              False         False
603   False            False          False              False         False
604   False            False          False              False         False
605   False            False          False              False         False
606   False            False          False              False         False

     company_location  company_size
0               False         False
1               False         False
2               False         False
3               False         False
4               False         False
..                ...           ...
602             False         False
603             False         False
604             False         False
605             False         False
606             False         False

[607 rows x 12 columns]
Unnamed: 0             int64
work_year              int64
experience_level      object
employment_type       object
job_title             object
salary                 int64
salary_currency       object
salary_in_usd          int64
employee_residence    object
remote_ratio           int64
company_location      object
company_size          object
dtype: object
```

Out[673…
```
RangeIndex(start=0, stop=607, step=1)
```

Question #2. compute average salary by year?

In [674…
```
mean_salary = df["salary_in_usd"].mean()
print("the average salary(USD) is:",mean_salary)
```

```
the average salary(USD) is: 112297.86985172982
```

Question#3. which year was the lowest and which year was the highest?

In [675…
```
# print(df["salary_in_usd"].argmax())
#data1= df.loc[(df["work_year"].isin(["2020"]))]

work_year=df['work_year'].unique()
print(work_year,type(df['work_year'][0]))
year_2020=df.loc[(df["work_year"]==2020),:]["salary_in_usd"].sum()
year_2021=df.loc[(df["work_year"]==2021),:]["salary_in_usd"].sum()
```

```
year_2022=df.loc[(df["work_year"]==2022),:]["salary_in_usd"].sum()

print(year_2020,year_2021,year_2022)
print("the highest year is 2022,the value is :",max(year_2020,year_2021,year_202
print("the lowest year is 2020,the value is :",min(year_2020,year_2021,year_2022
```

```
[2020 2021 2022] <class 'numpy.int64'>
6898536 21668273 39597998
the highest year is 2022,the value is : 39597998
the lowest year is 2020,the value is : 6898536
```

Question#4. for each experience level, compute the average salary (over 3 years) for each job title?

In [676…

```
job_title=df["job_title"].unique()
exp_level=df["experience_level"].unique()
#print(exp_level)
jb_exp_salary={}
for i in job_title:
    for j in exp_level:
        tuple_tamp=(i,j)
        jb_exp_salary[tuple_tamp]=df.loc[(df["job_title"]==i) & (df["experience_
        #print(df.loc[(df["job_title"]==i),:]["salary_in_usd"].sum())
print(jb_exp_salary)
```

```
{('Data Scientist', 'MI'): 82039.13333333333, ('Data Scientist', 'SE'): 152971.0
163934426, ('Data Scientist', 'EN'): 55330.90909090909, ('Data Scientist', 'E
X'): nan, ('Machine Learning Scientist', 'MI'): 109325.0, ('Machine Learning Sci
entist', 'SE'): 201666.66666666666, ('Machine Learning Scientist', 'EN'): 22500
0.0, ('Machine Learning Scientist', 'EX'): nan, ('Big Data Engineer', 'MI'): 335
37.0, ('Big Data Engineer', 'SE'): 111535.5, ('Big Data Engineer', 'EN'): 30703.
333333333332, ('Big Data Engineer', 'EX'): nan, ('Product Data Analyst', 'MI'):
13036.0, ('Product Data Analyst', 'SE'): nan, ('Product Data Analyst', 'EN'): na
n, ('Product Data Analyst', 'EX'): nan, ('Machine Learning Engineer', 'MI'): 744
66.58333333333, ('Machine Learning Engineer', 'SE'): 131176.0, ('Machine Learnin
g Engineer', 'EN'): 86996.33333333333, ('Machine Learning Engineer', 'EX'): nan,
('Data Analyst', 'MI'): 71699.20689655172, ('Data Analyst', 'SE'): 111922.629629
62964, ('Data Analyst', 'EN'): 53960.666666666664, ('Data Analyst', 'EX'): 12000
0.0, ('Lead Data Scientist', 'MI'): 115000.0, ('Lead Data Scientist', 'SE'): 115
285.0, ('Lead Data Scientist', 'EN'): nan, ('Lead Data Scientist', 'EX'): nan,
('Business Data Analyst', 'MI'): 74784.66666666667, ('Business Data Analyst', 'S
E'): nan, ('Business Data Analyst', 'EN'): 79551.0, ('Business Data Analyst', 'E
X'): nan, ('Lead Data Engineer', 'MI'): 56000.0, ('Lead Data Engineer', 'SE'): 1
66040.0, ('Lead Data Engineer', 'EN'): nan, ('Lead Data Engineer', 'EX'): 11818
7.0, ('Lead Data Analyst', 'MI'): 53304.5, ('Lead Data Analyst', 'SE'): 170000.
0, ('Lead Data Analyst', 'EN'): nan, ('Lead Data Analyst', 'EX'): nan, ('Data En
gineer', 'MI'): 85985.6603773585, ('Data Engineer', 'SE'): 137035.84126984127,
('Data Engineer', 'EN'): 58933.5, ('Data Engineer', 'EX'): 245500.0, ('Data Scie
nce Consultant', 'MI'): 103000.0, ('Data Science Consultant', 'SE'): nan, ('Data
Science Consultant', 'EN'): 62640.8, ('Data Science Consultant', 'EX'): 69741.0,
('BI Data Analyst', 'MI'): 78086.33333333333, ('BI Data Analyst', 'SE'): nan,
('BI Data Analyst', 'EN'): 32136.0, ('BI Data Analyst', 'EX'): 150000.0, ('Direc
tor of Data Science', 'MI'): nan, ('Director of Data Science', 'SE'): 168000.0,
('Director of Data Science', 'EN'): nan, ('Director of Data Science', 'EX'): 199
586.33333333334, ('Research Scientist', 'MI'): 136498.14285714287, ('Research Sc
ientist', 'SE'): 88859.4, ('Research Scientist', 'EN'): 86132.0, ('Research Scie
ntist', 'EX'): nan, ('Machine Learning Manager', 'MI'): nan, ('Machine Learning
```

Manager', 'SE'): 117104.0, ('Machine Learning Manager', 'EN'): nan, ('Machine Le
arning Manager', 'EX'): nan, ('Data Engineering Manager', 'MI'): 59303.0, ('Data
Engineering Manager', 'SE'): 159000.0, ('Data Engineering Manager', 'EN'): nan,
('Data Engineering Manager', 'EX'): 79833.0, ('Machine Learning Infrastructure E
ngineer', 'MI'): 54217.5, ('Machine Learning Infrastructure Engineer', 'SE'): 19
5000.0, ('Machine Learning Infrastructure Engineer', 'EN'): nan, ('Machine Learn
ing Infrastructure Engineer', 'EX'): nan, ('ML Engineer', 'MI'): 137025.0, ('ML
Engineer', 'SE'): 256000.0, ('ML Engineer', 'EN'): 18974.5, ('ML Engineer', 'E
X'): nan, ('AI Scientist', 'MI'): 160000.0, ('AI Scientist', 'SE'): 55000.0, ('A
I Scientist', 'EN'): 21987.25, ('AI Scientist', 'EX'): nan, ('Computer Vision En
gineer', 'MI'): nan, ('Computer Vision Engineer', 'SE'): 34302.333333333336, ('C
omputer Vision Engineer', 'EN'): 54536.333333333336, ('Computer Vision Enginee
r', 'EX'): nan, ('Principal Data Scientist', 'MI'): 151000.0, ('Principal Data S
cientist', 'SE'): 187939.4, ('Principal Data Scientist', 'EN'): nan, ('Principal
Data Scientist', 'EX'): 416000.0, ('Data Science Manager', 'MI'): 200000.0, ('Da
ta Science Manager', 'SE'): 149994.2, ('Data Science Manager', 'EN'): nan, ('Dat
a Science Manager', 'EX'): nan, ('Head of Data', 'MI'): 32974.0, ('Head of Dat
a', 'SE'): 151419.5, ('Head of Data', 'EN'): nan, ('Head of Data', 'EX'): 23250
0.0, ('3D Computer Vision Researcher', 'MI'): 5409.0, ('3D Computer Vision Resea
rcher', 'SE'): nan, ('3D Computer Vision Researcher', 'EN'): nan, ('3D Computer
Vision Researcher', 'EX'): nan, ('Data Analytics Engineer', 'MI'): 110000.0, ('D
ata Analytics Engineer', 'SE'): 64598.5, ('Data Analytics Engineer', 'EN'): 2000
0.0, ('Data Analytics Engineer', 'EX'): nan, ('Applied Data Scientist', 'MI'): 1
05619.0, ('Applied Data Scientist', 'SE'): 278500.0, ('Applied Data Scientist',
'EN'): 110037.0, ('Applied Data Scientist', 'EX'): nan, ('Marketing Data Analys
t', 'MI'): nan, ('Marketing Data Analyst', 'SE'): 88654.0, ('Marketing Data Anal
yst', 'EN'): nan, ('Marketing Data Analyst', 'EX'): nan, ('Cloud Data Engineer',
'MI'): 89294.0, ('Cloud Data Engineer', 'SE'): 160000.0, ('Cloud Data Engineer',
'EN'): nan, ('Cloud Data Engineer', 'EX'): nan, ('Financial Data Analyst', 'M
I'): 450000.0, ('Financial Data Analyst', 'SE'): nan, ('Financial Data Analyst',
'EN'): 100000.0, ('Financial Data Analyst', 'EX'): nan, ('Computer Vision Softwa
re Engineer', 'MI'): 95746.0, ('Computer Vision Software Engineer', 'SE'): nan,
('Computer Vision Software Engineer', 'EN'): 110000.0, ('Computer Vision Softwar
e Engineer', 'EX'): nan, ('Director of Data Engineering', 'MI'): nan, ('Director
of Data Engineering', 'SE'): 156738.0, ('Director of Data Engineering', 'EN'): n
an, ('Director of Data Engineering', 'EX'): nan, ('Data Science Engineer', 'M
I'): 40189.0, ('Data Science Engineer', 'SE'): 93610.5, ('Data Science Enginee
r', 'EN'): nan, ('Data Science Engineer', 'EX'): nan, ('Principal Data Enginee
r', 'MI'): nan, ('Principal Data Engineer', 'SE'): 192500.0, ('Principal Data En
gineer', 'EN'): nan, ('Principal Data Engineer', 'EX'): 600000.0, ('Machine Lear
ning Developer', 'MI'): 78791.0, ('Machine Learning Developer', 'SE'): 78791.0,
('Machine Learning Developer', 'EN'): 100000.0, ('Machine Learning Developer',
'EX'): nan, ('Applied Machine Learning Scientist', 'MI'): 178800.0, ('Applied Ma
chine Learning Scientist', 'SE'): nan, ('Applied Machine Learning Scientist', 'E
N'): 31875.0, ('Applied Machine Learning Scientist', 'EX'): nan, ('Data Analytic
s Manager', 'MI'): nan, ('Data Analytics Manager', 'SE'): 127134.28571428571,
('Data Analytics Manager', 'EN'): nan, ('Data Analytics Manager', 'EX'): nan,
('Head of Data Science', 'MI'): 110000.0, ('Head of Data Science', 'SE'): nan,
('Head of Data Science', 'EN'): nan, ('Head of Data Science', 'EX'): 158958.3333
3333334, ('Data Specialist', 'MI'): nan, ('Data Specialist', 'SE'): 165000.0,
('Data Specialist', 'EN'): nan, ('Data Specialist', 'EX'): nan, ('Data Architec
t', 'MI'): 166666.66666666666, ('Data Architect', 'SE'): 182076.625, ('Data Arch
itect', 'EN'): nan, ('Data Architect', 'EX'): nan, ('Finance Data Analyst', 'M
I'): nan, ('Finance Data Analyst', 'SE'): 61896.0, ('Finance Data Analyst', 'E
N'): nan, ('Finance Data Analyst', 'EX'): nan, ('Principal Data Analyst', 'MI'):
75000.0, ('Principal Data Analyst', 'SE'): 170000.0, ('Principal Data Analyst',
'EN'): nan, ('Principal Data Analyst', 'EX'): nan, ('Big Data Architect', 'MI'):
nan, ('Big Data Architect', 'SE'): 99703.0, ('Big Data Architect', 'EN'): nan,
('Big Data Architect', 'EX'): nan, ('Staff Data Scientist', 'MI'): nan, ('Staff
Data Scientist', 'SE'): 105000.0, ('Staff Data Scientist', 'EN'): nan, ('Staff D

```
ata Scientist', 'EX'): nan, ('Analytics Engineer', 'MI'): nan, ('Analytics Engin
eer', 'SE'): 195000.0, ('Analytics Engineer', 'EN'): nan, ('Analytics Engineer',
'EX'): 155000.0, ('ETL Developer', 'MI'): 54957.0, ('ETL Developer', 'SE'): nan,
('ETL Developer', 'EN'): nan, ('ETL Developer', 'EX'): nan, ('Head of Machine Le
arning', 'MI'): nan, ('Head of Machine Learning', 'SE'): nan, ('Head of Machine
Learning', 'EN'): nan, ('Head of Machine Learning', 'EX'): 79039.0, ('NLP Engine
er', 'MI'): 37236.0, ('NLP Engineer', 'SE'): nan, ('NLP Engineer', 'EN'): nan,
('NLP Engineer', 'EX'): nan, ('Lead Machine Learning Engineer', 'MI'): nan, ('Le
ad Machine Learning Engineer', 'SE'): 87932.0, ('Lead Machine Learning Enginee
r', 'EN'): nan, ('Lead Machine Learning Engineer', 'EX'): nan, ('Data Analytics
Lead', 'MI'): nan, ('Data Analytics Lead', 'SE'): 405000.0, ('Data Analytics Lea
d', 'EN'): nan, ('Data Analytics Lead', 'EX'): nan}
```

Question#5. For which job title was it the minimum and maximum?

In [677…

```
jb_salary_avg={}
job_title_max_value=job_title_min_value=df["salary_in_usd"].mean()
job_title_max=job_title_min=""
for i in job_title:
    jb_salary_avg[i]=df.loc[(df["job_title"]==i) ,:]["salary_in_usd"].mean()
    if job_title_max_value < df.loc[(df["job_title"]==i) ,:]["salary_in_usd"].me
        job_title_max = i
        job_title_max_value = df.loc[(df["job_title"]==i) ,:]["salary_in_usd"].m
    elif job_title_min_value > df.loc[(df["job_title"]==i) ,:]["salary_in_usd"].
        job_title_min = i
        job_title_min_value = df.loc[(df["job_title"]==i) ,:]["salary_in_usd"].m
print("the maximun salary job title is:",job_title_max,job_title_max_value,)
print("the minimun salary job title is:",job_title_min,job_title_min_value)
```

```
the maximun salary job title is: Data Analytics Lead 405000.0
the minimun salary job title is: 3D Computer Vision Researcher 5409.0
```

Question#6. for each year and job title compute the annual salary?

In [678…

```
jb_annual_salary={}
for i in work_year:
    for j in job_title:
        tuple_tamp=(i,j)
        jb_annual_salary[tuple_tamp]=df.loc[(df["job_title"]==j) & (df["work_yea
print(jb_annual_salary)
```

```
{(2020, 'Data Scientist'): 1805381, (2020, 'Machine Learning Scientist'): 26000
0, (2020, 'Big Data Engineer'): 293071, (2020, 'Product Data Analyst'): 26072,
(2020, 'Machine Learning Engineer'): 626949, (2020, 'Data Analyst'): 318831, (20
20, 'Lead Data Scientist'): 305000, (2020, 'Business Data Analyst'): 235000, (20
20, 'Lead Data Engineer'): 181000, (2020, 'Lead Data Analyst'): 87000, (2020, 'D
ata Engineer'): 969782, (2020, 'Data Science Consultant'): 108707, (2020, 'BI Da
ta Analyst'): 98000, (2020, 'Director of Data Science'): 325000, (2020, 'Researc
h Scientist'): 492000, (2020, 'Machine Learning Manager'): 117104, (2020, 'Data
Engineering Manager'): 139136, (2020, 'Machine Learning Infrastructure Enginee
r'): 50180, (2020, 'ML Engineer'): 15966, (2020, 'AI Scientist'): 45896, (2020,
'Computer Vision Engineer'): 60000, (2020, 'Principal Data Scientist'): 148261,
```

(2020, 'Data Science Manager'): 190200, (2020, 'Head of Data'): 0, (2020, '3D Co
mputer Vision Researcher'): 0, (2020, 'Data Analytics Engineer'): 0, (2020, 'App
lied Data Scientist'): 0, (2020, 'Marketing Data Analyst'): 0, (2020, 'Cloud Dat
a Engineer'): 0, (2020, 'Financial Data Analyst'): 0, (2020, 'Computer Vision So
ftware Engineer'): 0, (2020, 'Director of Data Engineering'): 0, (2020, 'Data Sc
ience Engineer'): 0, (2020, 'Principal Data Engineer'): 0, (2020, 'Machine Learn
ing Developer'): 0, (2020, 'Applied Machine Learning Scientist'): 0, (2020, 'Dat
a Analytics Manager'): 0, (2020, 'Head of Data Science'): 0, (2020, 'Data Specia
list'): 0, (2020, 'Data Architect'): 0, (2020, 'Finance Data Analyst'): 0, (202
0, 'Principal Data Analyst'): 0, (2020, 'Big Data Architect'): 0, (2020, 'Staff
Data Scientist'): 0, (2020, 'Analytics Engineer'): 0, (2020, 'ETL Developer'):
0, (2020, 'Head of Machine Learning'): 0, (2020, 'NLP Engineer'): 0, (2020, 'Lea
d Machine Learning Engineer'): 0, (2020, 'Data Analytics Lead'): 0, (2021, 'Data
Scientist'): 3180228, (2021, 'Machine Learning Scientist'): 582000, (2021, 'Big
Data Engineer'): 122721, (2021, 'Product Data Analyst'): 0, (2021, 'Machine Lear
ning Engineer'): 1343002, (2021, 'Data Analyst'): 1351592, (2021, 'Lead Data Sci
entist'): 40570, (2021, 'Business Data Analyst'): 59102, (2021, 'Lead Data Engin
eer'): 539160, (2021, 'Lead Data Analyst'): 189609, (2021, 'Data Engineer'): 266
2481, (2021, 'Data Science Consultant'): 377238, (2021, 'BI Data Analyst'): 3505
31, (2021, 'Director of Data Science'): 843539, (2021, 'Research Scientist'): 83
0036, (2021, 'Machine Learning Manager'): 0, (2021, 'Data Engineering Manager'):
477000, (2021, 'Machine Learning Infrastructure Engineer'): 195000, (2021, 'ML E
ngineer'): 667075, (2021, 'AI Scientist'): 97053, (2021, 'Computer Vision Engine
er'): 71516, (2021, 'Principal Data Scientist'): 1195762, (2021, 'Data Science M
anager'): 858759, (2021, 'Head of Data'): 567839, (2021, '3D Computer Vision Res
earcher'): 5409, (2021, 'Data Analytics Engineer'): 239197, (2021, 'Applied Data
Scientist'): 164275, (2021, 'Marketing Data Analyst'): 88654, (2021, 'Cloud Data
Engineer'): 249294, (2021, 'Financial Data Analyst'): 450000, (2021, 'Computer V
ision Software Engineer'): 165746, (2021, 'Director of Data Engineering'): 31347
6, (2021, 'Data Science Engineer'): 167410, (2021, 'Principal Data Engineer'): 9
85000, (2021, 'Machine Learning Developer'): 100000, (2021, 'Applied Machine Lea
rning Scientist'): 461400, (2021, 'Data Analytics Manager'): 380000, (2021, 'Hea
d of Data Science'): 195000, (2021, 'Data Specialist'): 165000, (2021, 'Data Arc
hitect'): 500000, (2021, 'Finance Data Analyst'): 61896, (2021, 'Principal Data
Analyst'): 170000, (2021, 'Big Data Architect'): 99703, (2021, 'Staff Data Scien
tist'): 105000, (2021, 'Analytics Engineer'): 0, (2021, 'ETL Developer'): 0, (20
21, 'Head of Machine Learning'): 0, (2021, 'NLP Engineer'): 0, (2021, 'Lead Mach
ine Learning Engineer'): 0, (2021, 'Data Analytics Lead'): 0, (2022, 'Data Scien
tist'): 10485251, (2022, 'Machine Learning Scientist'): 425300, (2022, 'Big Data
Engineer'): 0, (2022, 'Product Data Analyst'): 0, (2022, 'Machine Learning Engin
eer'): 2330135, (2022, 'Data Analyst'): 7340204, (2022, 'Lead Data Scientist'):
0, (2022, 'Business Data Analyst'): 89354, (2022, 'Lead Data Engineer'): 118187,
(2022, 'Lead Data Analyst'): 0, (2022, 'Data Engineer'): 11247437, (2022, 'Data
Science Consultant'): 0, (2022, 'BI Data Analyst'): 0, (2022, 'Director of Data
Science'): 196979, (2022, 'Research Scientist'): 422276, (2022, 'Machine Learnin
g Manager'): 0, (2022, 'Data Engineering Manager'): 0, (2022, 'Machine Learning
Infrastructure Engineer'): 58255, (2022, 'ML Engineer'): 21983, (2022, 'AI Scien
tist'): 320000, (2022, 'Computer Vision Engineer'): 135000, (2022, 'Principal Da
ta Scientist'): 162674, (2022, 'Data Science Manager'): 850983, (2022, 'Head of
Data'): 232974, (2022, '3D Computer Vision Researcher'): 0, (2022, 'Data Analyti
cs Engineer'): 20000, (2022, 'Applied Data Scientist'): 714000, (2022, 'Marketin
g Data Analyst'): 0, (2022, 'Cloud Data Engineer'): 0, (2022, 'Financial Data An
alyst'): 100000, (2022, 'Computer Vision Software Engineer'): 150000, (2022, 'Di
rector of Data Engineering'): 0, (2022, 'Data Science Engineer'): 60000, (2022,
'Principal Data Engineer'): 0, (2022, 'Machine Learning Developer'): 157582, (20
22, 'Applied Machine Learning Scientist'): 106875, (2022, 'Data Analytics Manage
r'): 509940, (2022, 'Head of Data Science'): 391875, (2022, 'Data Specialist'):
0, (2022, 'Data Architect'): 1456613, (2022, 'Finance Data Analyst'): 0, (2022,
'Principal Data Analyst'): 75000, (2022, 'Big Data Architect'): 0, (2022, 'Staff
Data Scientist'): 0, (2022, 'Analytics Engineer'): 700000, (2022, 'ETL Develope

```
r'): 109914, (2022, 'Head of Machine Learning'): 79039, (2022, 'NLP Engineer'):
37236, (2022, 'Lead Machine Learning Engineer'): 87932, (2022, 'Data Analytics L
ead'): 405000}
```

Question#7. which job title experienced the maximum change in salary? Which job title had the minimum change.

In [679…

```
jb_salary_gap={}
for i in job_title:
    jb_salary_gap[i]=df.loc[(df["job_title"]==i), : ]["salary_in_usd"].max() - d

print(jb_salary_gap.values())
jb_salary_gap_max=jb_salary_gap_min=""
jb_salary_gap_max_val=jb_salary_gap_min_val=188000
for i in job_title:
    if jb_salary_gap_max_val < jb_salary_gap[i]:
        jb_salary_gap_max_val = jb_salary_gap[i]
        jb_salary_gap_max = i
    elif jb_salary_gap_min_val > jb_salary_gap[i] & jb_salary_gap[i]!= 0 :
        jb_salary_gap_min_val = jb_salary_gap[i]
        jb_salary_gap_min = i
print("the maximun change job title is ",jb_salary_gap_max,", the value is ",jb_
print("the minimun change jon title is ",jb_salary_gap_min,", the value is ",jb_
```

```
dict_values([409141, 248000, 108165, 13928, 230000, 193928, 149430, 116558, 2200
00, 150391, 320000, 97293, 140728, 194974, 408000, 0, 114697, 144820, 254034, 18
8000, 115000, 267739, 186906, 202026, 0, 90000, 325762, 0, 70706, 350000, 80000,
86524, 87032, 415000, 21209, 391125, 44860, 139000, 0, 175700, 0, 95000, 0, 0, 7
0300, 0, 0, 0, 0, 0])
the maximun change job title is  Principal Data Engineer , the value is  415000
the minimun change jon title is  Product Data Analyst , the value is  13928
```

Question#8. compute the average salaries for each remote ratio? how many entries are there?

In [680…

```
remote_ratio=df["remote_ratio"].unique()
rr_avg_dict={}
print(remote_ratio)
for i in remote_ratio:
    rr_avg_dict[i]=df.loc[(df["remote_ratio"]==i) ,:]["salary_in_usd"].mean()
print(rr_avg_dict)
```

```
[  0  50 100]
{0: 106354.62204724409, 50: 80823.0303030303, 100: 122457.45406824147}
```

Question#9. which company locations pay the highest salary and lowest salaries.

In [681…

```
location_salary=0
```

```
employee_residence=df["employee_residence"].unique()
employee_residence_dict={}
print(employee_residence)
for i in employee_residence:
    employee_residence_dict[i]=df.loc[(df["employee_residence"]==i), : ]["salary
#print(employee_residence_dict)
employee_residence_min=employee_residence_max=""
employee_residence_min_val=employee_residence_max_val=76738.66666666667

for i in employee_residence:
    if employee_residence_min_val > employee_residence_dict[i] :
        employee_residence_min_val=employee_residence_dict[i]
        employee_residence_min = i
    elif employee_residence_max_val <  employee_residence_dict[i]:
        employee_residence_max_val=employee_residence_dict[i]
        employee_residence_max = i
print("the lowest place is ",employee_residence_min,", value is ",employee_resid
print("the lowest place is ",employee_residence_max,", value is ",employee_resid
```

```
['DE' 'JP' 'GB' 'HN' 'US' 'HU' 'NZ' 'FR' 'IN' 'PK' 'PL' 'PT' 'CN' 'GR'
 'AE' 'NL' 'MX' 'CA' 'AT' 'NG' 'PH' 'ES' 'DK' 'RU' 'IT' 'HR' 'BG' 'SG'
 'BR' 'IQ' 'VN' 'BE' 'UA' 'MT' 'CL' 'RO' 'IR' 'CO' 'MD' 'KE' 'SI' 'HK'
 'TR' 'RS' 'PR' 'LU' 'JE' 'CZ' 'AR' 'DZ' 'TN' 'MY' 'EE' 'AU' 'BO' 'IE'
 'CH']
the lowest place is  IR , value is  4000.0
the lowest place is  MY , value is  200000.0
```

Question#10. Would you change your resume after analyzing this dataset? his data does give me a new understanding of the salary of CS majors, but I do not intend to change my resume. One of the big reasons is that I'm already hiring, with my existing knowledge base. Not enough for other jobs.