```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")
sns.set(style="whitegrid", color_codes=True)
pd.set_option('display.max_columns', None)

# Loading the dataset
data = pd.read_excel('/content/data.xlsx')
# checking the dimensions of dataset
data.shape
```

```
(3998, 39)
```

```
data.columns
```

```
Index(['Unnamed: 0', 'ID', 'Salary', 'DOJ', 'DOL', 'Designation',
'JobCity',
       'Gender', 'DOB', '10percentage', '10board', '12graduation',
       '12percentage', '12board', 'CollegeID', 'CollegeTier',
'Degree',
       'Specialization', 'collegeGPA', 'CollegeCityID',
'CollegeCityTier',
       'CollegeState', 'GraduationYear', 'English', 'Logical',
'Quant',
       'Domain', 'ComputerProgramming', 'ElectronicsAndSemicon',
       'ComputerScience', 'MechanicalEngg', 'ElectricalEngg',
'TelecomEngg',
       'CivilEngg', 'conscientiousness', 'agreeableness',
'extraversion',
       'nueroticism', 'openess_to_experience'],
      dtype='object')
```

```
dataset = data.iloc[:,1:]
print(f'Rows : {dataset.shape[0]}\nColumns:{dataset.shape[1]}')
dataset.info()
```

```
Rows : 3998
Columns:38
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3998 entries, 0 to 3997
Data columns (total 38 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   ID                     3998 non-null   int64
 1   Salary                 3998 non-null   int64
 2   DOJ                    3998 non-null   datetime64[ns]
 3   DOL                    3998 non-null   object
 4   Designation            3998 non-null   object
```

```
 5   JobCity                 3998 non-null   object
 6   Gender                  3998 non-null   object
 7   DOB                     3998 non-null   datetime64[ns]
 8   10percentage            3998 non-null   float64
 9   10board                 3998 non-null   object
 10  12graduation            3998 non-null   int64
 11  12percentage            3998 non-null   float64
 12  12board                 3998 non-null   object
 13  CollegeID               3998 non-null   int64
 14  CollegeTier             3998 non-null   int64
 15  Degree                  3998 non-null   object
 16  Specialization          3998 non-null   object
 17  collegeGPA              3998 non-null   float64
 18  CollegeCityID           3998 non-null   int64
 19  CollegeCityTier         3998 non-null   int64
 20  CollegeState            3998 non-null   object
 21  GraduationYear          3998 non-null   int64
 22  English                 3998 non-null   int64
 23  Logical                 3998 non-null   int64
 24  Quant                   3998 non-null   int64
 25  Domain                  3998 non-null   float64
 26  ComputerProgramming     3998 non-null   int64
 27  ElectronicsAndSemicon   3998 non-null   int64
 28  ComputerScience         3998 non-null   int64
 29  MechanicalEngg          3998 non-null   int64
 30  ElectricalEngg          3998 non-null   int64
 31  TelecomEngg             3998 non-null   int64
 32  CivilEngg               3998 non-null   int64
 33  conscientiousness       3998 non-null   float64
 34  agreeableness           3998 non-null   float64
 35  extraversion            3998 non-null   float64
 36  nueroticism             3998 non-null   float64
 37  openess_to_experience   3998 non-null   float64
dtypes: datetime64[ns](2), float64(9), int64(18), object(9)
memory usage: 1.2+ MB
```

```
dataset.head()
```

{"type":"dataframe","variable_name":"dataset"}

```
np.round(dataset.describe())
```

{"type":"dataframe"}

Data Cleaning

```
for i in dataset.columns:
 print('*'*20,i,'*'*20)
 print(dataset[i].unique())
```

```
******************** ID ********************
[203097 579905 810601 ... 355888 947111 324966]
******************** Salary ********************
[ 420000  500000  325000 1100000  200000  300000  400000  600000
230000
  450000  270000  350000  250000  120000  320000  190000  180000
335000
  435000  345000  145000  220000  530000  340000  360000  215000
80000
  330000  380000  110000  205000   95000  390000   60000  240000
525000
  305000  150000  310000  455000  800000  100000  280000  445000
315000
  370000  275000 1500000  425000  470000  460000  510000  480000
170000
  640000  225000  440000 1200000  675000  105000  195000  385000
235000
  615000  290000  140000  405000 1860000  375000  430000  660000
70000
  410000  550000   35000  115000  415000  265000  285000  245000
395000
  560000  700000  185000  160000  625000   85000  135000  785000
210000
  155000  355000  535000  690000  260000 1110000 1000000  505000
475000
  715000  820000   90000  720000 2600000  515000   55000  495000
65000
  655000  545000  520000  645000 1025000  775000  490000 1300000
3500000
  910000  570000  255000  130000  175000  730000  555000  465000
680000
  165000  630000  365000 1050000 2000000  860000  125000   50000
580000
  485000 4000000 2020000  650000   45000  610000  760000  585000
620000
  870000 2050000  540000  144000  605000 1320000  755000  880000
3000000
   75000  295000   40000  575000  565000 2500000 2300000  590000
950000
 1800000  725000  930000  750000  705000 1745000  850000  845000
670000
 1030000  770000  900000 1210000  810000  925000]
******************** DOJ ********************
```

```
['2012-06-01T00:00:00.000000000' '2013-09-01T00:00:00.000000000'
 '2014-06-01T00:00:00.000000000' '2011-07-01T00:00:00.000000000'
 '2014-03-01T00:00:00.000000000' '2014-08-01T00:00:00.000000000'
 '2014-07-01T00:00:00.000000000' '2013-07-01T00:00:00.000000000'
 '2011-04-01T00:00:00.000000000' '2011-08-01T00:00:00.000000000'
 '2013-12-01T00:00:00.000000000' '2014-01-01T00:00:00.000000000'
 '2013-08-01T00:00:00.000000000' '2014-09-01T00:00:00.000000000'
 '2010-11-01T00:00:00.000000000' '2012-08-01T00:00:00.000000000'
 '2013-10-01T00:00:00.000000000' '2012-09-01T00:00:00.000000000'
 '2011-01-01T00:00:00.000000000' '2015-02-01T00:00:00.000000000'
 '2014-11-01T00:00:00.000000000' '2011-12-01T00:00:00.000000000'
 '2014-10-01T00:00:00.000000000' '2015-01-01T00:00:00.000000000'
 '2013-03-01T00:00:00.000000000' '2010-10-01T00:00:00.000000000'
 '2013-01-01T00:00:00.000000000' '2011-06-01T00:00:00.000000000'
 '2014-04-01T00:00:00.000000000' '2012-05-01T00:00:00.000000000'
 '2012-10-01T00:00:00.000000000' '2015-04-01T00:00:00.000000000'
 '2012-03-01T00:00:00.000000000' '2013-06-01T00:00:00.000000000'
 '2009-09-01T00:00:00.000000000' '2013-11-01T00:00:00.000000000'
 '2010-07-01T00:00:00.000000000' '2014-02-01T00:00:00.000000000'
 '2015-06-01T00:00:00.000000000' '2014-05-01T00:00:00.000000000'
 '2014-12-01T00:00:00.000000000' '2011-11-01T00:00:00.000000000'
 '2015-07-01T00:00:00.000000000' '2013-05-01T00:00:00.000000000'
 '2011-03-01T00:00:00.000000000' '2015-03-01T00:00:00.000000000'
 '2012-07-01T00:00:00.000000000' '2011-10-01T00:00:00.000000000'
 '2010-04-01T00:00:00.000000000' '2013-04-01T00:00:00.000000000'
 '2010-12-01T00:00:00.000000000' '2013-02-01T00:00:00.000000000'
 '2011-09-01T00:00:00.000000000' '2012-02-01T00:00:00.000000000'
 '2012-01-01T00:00:00.000000000' '2012-12-01T00:00:00.000000000'
 '2010-09-01T00:00:00.000000000' '2012-04-01T00:00:00.000000000'
 '2012-11-01T00:00:00.000000000' '2015-05-01T00:00:00.000000000'
 '2010-06-01T00:00:00.000000000' '2011-02-01T00:00:00.000000000'
 '2010-08-01T00:00:00.000000000' '2010-05-01T00:00:00.000000000'
 '2011-05-01T00:00:00.000000000' '2004-08-01T00:00:00.000000000'
 '2008-11-01T00:00:00.000000000' '2009-06-01T00:00:00.000000000'
 '2010-02-01T00:00:00.000000000' '2009-11-01T00:00:00.000000000'
 '2010-03-01T00:00:00.000000000' '2015-11-01T00:00:00.000000000'
 '2006-01-01T00:00:00.000000000' '2015-08-01T00:00:00.000000000'
 '2010-01-01T00:00:00.000000000' '2015-12-01T00:00:00.000000000'
 '2007-09-01T00:00:00.000000000' '1991-06-01T00:00:00.000000000'
 '2007-07-01T00:00:00.000000000' '2007-06-01T00:00:00.000000000'
 '2007-02-01T00:00:00.000000000']
********** DOL **********
['present' datetime.datetime(2015, 3, 1, 0, 0)
 datetime.datetime(2015, 5, 1, 0, 0) datetime.datetime(2015, 7, 1, 0,
0)
 datetime.datetime(2015, 4, 1, 0, 0) datetime.datetime(2014, 10, 1, 0,
0)
 datetime.datetime(2014, 9, 1, 0, 0) datetime.datetime(2014, 6, 1, 0,
0)
 datetime.datetime(2012, 9, 1, 0, 0) datetime.datetime(2013, 12, 1, 0,
```

```
0)
 datetime.datetime(2015, 6, 1, 0, 0) datetime.datetime(2013, 10, 1, 0,
0)
 datetime.datetime(2015, 1, 1, 0, 0) datetime.datetime(2014, 4, 1, 0,
0)
 datetime.datetime(2013, 6, 1, 0, 0) datetime.datetime(2012, 3, 1, 0,
0)
 datetime.datetime(2014, 7, 1, 0, 0) datetime.datetime(2013, 2, 1, 0,
0)
 datetime.datetime(2014, 1, 1, 0, 0) datetime.datetime(2013, 4, 1, 0,
0)
 datetime.datetime(2012, 7, 1, 0, 0) datetime.datetime(2014, 5, 1, 0,
0)
 datetime.datetime(2013, 9, 1, 0, 0) datetime.datetime(2015, 2, 1, 0,
0)
 datetime.datetime(2012, 1, 1, 0, 0) datetime.datetime(2015, 8, 1, 0,
0)
 datetime.datetime(2014, 8, 1, 0, 0) datetime.datetime(2015, 12, 1, 0,
0)
 datetime.datetime(2014, 12, 1, 0, 0) datetime.datetime(2012, 5, 1, 0,
0)
 datetime.datetime(2011, 3, 1, 0, 0) datetime.datetime(2011, 7, 1, 0,
0)
 datetime.datetime(2014, 2, 1, 0, 0) datetime.datetime(2011, 12, 1, 0,
0)
 datetime.datetime(2015, 10, 1, 0, 0) datetime.datetime(2014, 11, 1,
0, 0)
 datetime.datetime(2014, 3, 1, 0, 0) datetime.datetime(2011, 11, 1, 0,
0)
 datetime.datetime(2013, 5, 1, 0, 0) datetime.datetime(2013, 7, 1, 0,
0)
 datetime.datetime(2013, 11, 1, 0, 0) datetime.datetime(2011, 1, 1, 0,
0)
 datetime.datetime(2011, 5, 1, 0, 0) datetime.datetime(2012, 2, 1, 0,
0)
 datetime.datetime(2012, 11, 1, 0, 0) datetime.datetime(2012, 6, 1, 0,
0)
 datetime.datetime(2013, 8, 1, 0, 0) datetime.datetime(2005, 3, 1, 0,
0)
 datetime.datetime(2013, 3, 1, 0, 0) datetime.datetime(2012, 10, 1, 0,
0)
 datetime.datetime(2011, 2, 1, 0, 0) datetime.datetime(2010, 2, 1, 0,
0)
 datetime.datetime(2013, 1, 1, 0, 0) datetime.datetime(2011, 6, 1, 0,
0)
 datetime.datetime(2015, 9, 1, 0, 0) datetime.datetime(2012, 4, 1, 0,
0)
 datetime.datetime(2012, 8, 1, 0, 0) datetime.datetime(2011, 4, 1, 0,
0)
 datetime.datetime(2011, 10, 1, 0, 0) datetime.datetime(2015, 11, 1,
```

```
0, 0)
 datetime.datetime(2012, 12, 1, 0, 0) datetime.datetime(2011, 9, 1, 0,
0)
 datetime.datetime(2010, 8, 1, 0, 0) datetime.datetime(2011, 8, 1, 0,
0)
 datetime.datetime(2009, 6, 1, 0, 0) datetime.datetime(2008, 3, 1, 0,
0)
 datetime.datetime(2010, 10, 1, 0, 0)]
********** Designation **********
['senior quality engineer' 'assistant manager' 'systems engineer'
 'senior software engineer' 'get' 'system engineer'
 'java software engineer' 'mechanical engineer' 'electrical engineer'
 'project engineer' 'senior php developer' 'senior systems engineer'
 'quality assurance engineer' 'qa analyst' 'network engineer'
 'product development engineer' 'associate software developer'
 'data entry operator' 'software engineer' 'developer'
 'electrical project engineer' 'programmer analyst' 'systems analyst'
 'ase' 'telecommunication engineer' 'application developer'
 'ios developer' 'executive assistant' 'online marketing manager'
 'documentation specialist' 'associate software engineer'
 'management trainee' 'site manager' 'software developer' '.net
developer'
 'production engineer' 'jr. software engineer'
 'trainee software developer' 'ui developer' 'assistant system
engineer'
 'android developer' 'customer service' 'test engineer' 'java
developer'
 'engineer' 'recruitment coordinator' 'technical support engineer'
 'data analyst' 'assistant software engineer' 'faculty'
 'entry level management trainee' 'customer service representative'
 'software test engineer' 'firmware engineer' 'php developer'
 'research associate' 'research analyst' 'quality engineer'
'programmer'
 'technical support executive' 'business analyst' 'web developer'
 'application engineer' 'project coordinator' 'engineer trainee'
 'sap consultant' 'quality analyst' 'marketing coordinator'
 'system administrator' 'senior engineer' 'business development
managerde'
 'network administrator' 'technical support specialist'
 'business development executive' 'junior software engineer'
 'asp.net developer' 'graduate engineer trainee' 'field engineer'
 'assistant professor' 'trainee software engineer'
 'senior software developer' 'quality assurance automation engineer'
 'design engineer' 'telecom engineer' 'quality control engineer'
 'hardware engineer' 'hr recruiter' 'sales associate' 'junior
engineer'
 'associate engineer' 'maintenance engineer' 'sales engineer'
 'human resources associate' 'mobile application developer'
 'electronic field service engineer' 'process associate'
 'field service engineer' 'it support specialist'
```

```
 'software development engineer' 'business process analyst'
 'operation engineer' 'electrical designer' 'marketing assistant'
 'sales executive' 'admin assistant' 'senior java developer'
 'account executive' 'oracle dba' 'rf engineer'
 'embedded software engineer' 'programmer analyst trainee'
 'technical engineer' 'operations executive' 'trainee engineer'
 'recruiter' 'lecturer' '.net web developer' 'marketing executive'
 'operations assistant' 'associate manager' 'electrical design
engineer'
 'systems administrator' 'client services associate' 'it analyst'
 'senior developer' 'cad designer' 'business technology analyst'
 'asst. manager' 'service engineer' 'executive recruiter'
 'planning engineer' 'associate technical operations' 'web designer'
 'software architect' 'software quality assurance tester' 'seo
trainee'
 'process engineer' 'software quality assurance analyst' 'designer'
 'business systems consultant' 'business development manager'
 'junior research fellow' 'technical recruiter' 'operations analyst'
 'quality assurance test engineer' 'linux systems administrator'
 'software trainee' 'entry level sales and marketing'
 'electrical field engineer' 'windows systems administrator'
 'junior software developer' 'python developer'
 'web application developer' 'assistant systems engineer'
 'javascript developer' 'operation executive' 'performance engineer'
 'technical writer' 'operations engineer and jetty handling'
 'lead engineer' 'portfolio analyst' 'associate system engineer'
 'mechanical design engineer' 'product engineer'
 'network security engineer' 'operations manager' 'technical lead'
 'operations' 'quality assurance tester' 'automation engineer'
 'data scientist' 'quality associate' 'manual tester' 'sr. engineer'
 'embedded engineer' 'service and sales engineer'
 'telecom support engineer' 'engineer- customer support' 'cloud
engineer'
 'branch manager' 'business analyst consultant' 'technology lead'
 'software trainee engineer' 'dcs engineer' 'junior manager' 'ux
designer'
 'clerical' 'hr generalist' 'database administrator'
 'senior design engineer' 'seo' 'assistant engineer' 'marketing
analyst'
 'it executive' 'salesforce developer' 'software tester' 'sql dba'
 'junior engineer product support' 'manager' 'senior business analyst'
 'c# developer' 'implementation engineer' 'executive hr'
 'executive engineer' 'sharepoint developer' 'system analyst'
 'sales management trainee' 'senior project engineer' 'it recruiter'
 'software engineer analyst' 'desktop support technician'
 'continuous improvement engineer' 'process advisor' 'etl developer'
 'sales and service engineer' 'project manager' 'training specialist'
 'product manager' 'staffing recruiter' 'assistant programmer'
 'quality controller' 'mis executive' 'game developer'
 'digital marketing specialist' 'principal software engineer'
```

```
 'software devloper' 'senior mechanical engineer'
 'technical operations analyst' 'service coordinator' 'testing
engineer'
 'technical assistant' 'sap abap consultant' 'seo engineer'
 'project assistant' 'talent acquisition specialist'
 'sales account manager' 'software engineer trainee'
 'customer service manager' 'help desk analyst' 'general manager'
 'engineering manager' 'senior network engineer'
 'field based employee relations manager' 'phone banking officer'
 'support engineer' 'associate test engineer' 'technology analyst'
 'network support engineer' 'it business analyst' 'junior system
analyst'
 'senior .net developer' 'secretary' 'research engineer'
 'quality assurance auditor' 'process executive'
 'lecturer & electrical maintenance' 'office coordinator' 'hr manager'
 'html developer' 'sales support' 'front end web developer'
 'administrative support' 'territory sales manager'
 'project administrator' 'environmental engineer' 'web designer and
seo'
 'information security analyst' 'field business development associate'
 'operational executive' 'administrative coordinator'
 'senior risk consultant' 'desktop support engineer' 'cad drafter'
 'noc engineer' 'industrial engineer' 'it engineer'
 'human resources intern' 'senior quality assurance engineer'
 'clerical assistant' 'software enginner' 'quality assurance'
 'delivery software engineer' 'graphic designer'
 'sales development manager' 'visiting faculty'
 'business intelligence analyst' 'team lead'
 'operational excellence manager' 'sales & service engineer' 'web
intern'
 'full stack developer' 'database developer' 'sr. database engineer'
 'graduate apprentice trainee' 'software engineer associate'
 'technical analyst' 'executive engg' 'it technician'
 'business system analyst' 'process control engineer'
 'technical consultant' 'business office manager'
 'quality control inspector' 'product design engineer'
 'manufacturing engineer' 'seo executive' 'sap analyst'
 'software engineere' 'financial service consultant' 'co faculty'
 'software analyst' 'desktop support analyst' 'graduate engineer'
 'engineering technician' 'it assistant' 'marketing manager'
 'human resource assistant' 'hr assistant' 'product developer'
 'customer support engineer' 'quality control inspection technician'
 'gis/cad engineer' 'senior web developer' 'sql developer'
 'research staff member' 'sap abap associate consultant' 'associate
qa'
 'corporate recruiter' 'project management officer'
 'business systems analyst' 'software programmer' 'help desk
technician'
 'sales manager' 'catalog associate' 'assistant store manager'
 'software engg' 'it developer' 'apprentice' 'business consultant'
```

```
 'controls engineer' 'ruby on rails developer' 'risk consultant'
 'account manager' 'professor' 'assistant administrator' 'civil
engineer'
 'educator' 'service manager' 'teradata dba'
 'full-time loss prevention associate' 'junior recruiter'
 'associate developer' 'assistant electrical engineer' 'shift
engineer'
 'dotnet developer' 'rf/dt engineer' 'human resources analyst'
 'software test engineerte' 'junior .net developer' 'java trainee'
 'maintenance supervisor' 'r&d engineer' 'front end developer'
 'engineer-hws' 'operations engineer' 'senior research fellow'
 'web designer and joomla administrator' 'enterprise solutions
developer'
 'information technology specialist' 'site engineer'
 'graduate trainee engineer' 'quality assurance analyst' 'cnc
programmer'
 'financial analyst' 'system engineer trainee' 'sap mm consultant'
 'assistant system engineer trainee' 'qa trainee' 'teradata developer'
 'hr executive' 'senior programmer' 'software test engineer (etl)'
 'associate software engg' 'supply chain analyst' 'sales trainer'
 'software executive' 'team leader' 'assistant system engineer -
trainee'
 'seo analyst' 'risk investigator' 'executive administrative
assistant'
 'program manager' 'r & d' 'sap functional consultant'
 'website developer/tester' 'software designer' 'sales coordinator'
 'qa engineer' 'aircraft technician' 'customer care executive'
 'senior test engineer' 'program analyst trainee'
 'electrical controls engineer' 'trainee decision scientist' 'editor'
 'bss engineer' 'dba' 'software eng' 'computer faculty'
 'recruitment associate' 'logistics executive' 'quality consultant'
 'senior sales executive' 'db2 dba' 'test technician'
 'it operations associate' 'software engineering associate'
 'research scientist' 'jr. software developer']
******************** JobCity ********************
['Bangalore' 'Indore' 'Chennai' 'Gurgaon' 'Manesar' 'Hyderabad'
'Banglore'
 'Noida' 'Kolkata' 'Pune' -1 'mohali' 'Jhansi' 'Delhi' 'Hyderabad '
 'Bangalore ' 'noida' 'delhi' 'Bhubaneswar' 'Navi Mumbai' 'Mumbai'
 'New Delhi' 'Mangalore' 'Rewari' 'Gaziabaad' 'Bhiwadi' 'Mysore'
'Rajkot'
 'Greater Noida' 'Jaipur' 'noida ' 'HYDERABAD' 'mysore' 'THANE'
 'Maharajganj' 'Thiruvananthapuram' 'Punchkula' 'Bhubaneshwar' 'Pune '
 'coimbatore' 'Dhanbad' 'Lucknow' 'Trivandrum' 'kolkata' 'mumbai'
 'Gandhi Nagar' 'Una' 'Daman and Diu' 'chennai' 'GURGOAN'
'vsakhapttnam'
 'pune' 'Nagpur' 'Bhagalpur' 'new delhi - jaisalmer' 'Coimbatore'
 'Ahmedabad' 'Kochi/Cochin' 'Bankura' 'Bengaluru' 'Mysore ' 'Kanpur '
 'jaipur' 'Gurgaon ' 'bangalore' 'CHENNAI' 'Vijayawada' 'Kochi'
'Beawar'
```

```
 'Alwar' 'NOIDA' 'Greater noida' 'Siliguri ' 'raipur' 'gurgaon'
'Bhopal'
 'Faridabad' 'Jodhpur' 'udaipur' 'Muzaffarpur' 'Kolkata``
'Bulandshahar'
 'Haridwar' 'Raigarh' 'Visakhapatnam' 'Jabalpur' 'hyderabad' 'Unnao'
 'KOLKATA' 'Thane' 'Aurangabad' 'Belgaum' 'gurgoan' 'Dehradun'
'Rudrapur'
 'Jamshedpur' 'vizag' 'Nouda' 'Dharamshala' 'Banagalore' 'Hissar'
'Ranchi'
 'BANGALORE' 'Madurai' 'Gurga' 'Chandigarh' 'Australia' ' Chennai'
 'CHEYYAR' 'Mumbai ' 'sonepat' 'Ghaziabad' 'Pantnagar' 'Siliguri'
 'mumbai ' 'Jagdalpur' 'Chennai ' 'angul' 'Baroda' ' ariyalur' 'Jowai'
 'Kochi/Cochin, Chennai and Coimbatore' 'bhubaneswar' 'Neemrana'
'VIZAG'
 'Tirupathi' 'Lucknow ' 'Ahmedabad ' 'Bhubneshwar' 'Noida ' 'pune '
 'Calicut' 'Gandhinagar' 'LUCKNOW' 'Dubai' 'bengaluru' 'MUMBAI'
 'Ahmednagar' 'Nashik' 'New delhi' 'Bellary' 'Ludhiana' 'New Delhi '
 'Muzaffarnagar' 'BHOPAL' 'Gurgoan' 'Gagret' 'Indirapuram, Ghaziabad'
 'Gwalior' 'new delhi' 'TRIVANDRUM' 'Chennai & Mumbai' 'Rajasthan'
 'Sonipat' 'Bareli' 'Kanpur' 'Hospete' 'Miryalaguda' ' mumbai'
'Dharuhera'
 'lucknow' 'meerut' 'dehradun' 'Ganjam' 'Hubli' 'bangalore ' 'NAVI
MUMBAI'
 'ncr' 'Agra' 'Trichy' 'kudankulam ,tarapur' 'Ongole' 'Sambalpur'
 'Pondicherry' 'Bundi' 'SADULPUR,RAJGARH,DISTT-CHURU,RAJASTHAN' 'AM'
 'Bikaner' 'Vadodara' 'BAngalore' 'india' 'Asansol' 'Tirunelvelli'
 'Ernakulam' 'DELHI' 'Bilaspur' 'Chandrapur' 'Nanded' 'Dharmapuri'
 'Vandavasi' 'Rohtak' 'trivandrum' 'Nagpur ' 'Udaipur' 'Patna'
'banglore'
 'indore' 'Salem' 'Nasikcity' 'Gandhinagar ' 'Technopark, Trivandrum'
 'Bharuch' 'Tornagallu' 'Raipur' 'Kolkata ' 'Jaspur' 'Burdwan'
 'Bhubaneswar ' 'Shimla' 'ahmedabad' 'Gajiabaad' 'Jammu' 'Shahdol'
 'Muvattupuzha' 'Al Jubail,Saudi Arabia' 'Kalmar, Sweden'
'Secunderabad'
 'A-64,sec-64,noida' 'Ratnagiri' 'Jhajjar' 'Gulbarga'
 'hyderabad(bhadurpally)' 'Nalagarh' 'Chandigarh ' 'Jaipur '
 'Jeddah Saudi Arabia' ' Delhi' 'PATNA' 'SHAHDOL' 'Chennai, Bangalore'
 'Bhopal ' 'Jamnagar' 'PUNE' 'Tirupati' 'Gonda' 'jamnagar' 'chennai '
 'orissa' 'kharagpur' 'Trivandrum ' 'Navi Mumbai , Hyderabad'
'Joshimath'
 'chandigarh' 'Bathinda' 'Johannesburg' 'kala amb ' 'Karnal' 'LONDON'
 'Kota' 'Panchkula' 'Baddi HP' 'Nagari' 'Mettur, Tamil Nadu '
'Durgapur'
 'pondi' 'Surat' 'Kurnool' 'kolhapur' 'Madurai ' 'GREATER NOIDA'
'Bhilai'
 ' Pune' 'hderabad' 'KOTA' 'thane' 'Vizag' 'Bahadurgarh'
 'Rayagada, Odisha' 'kakinada' 'GURGAON' 'Varanasi' 'punr' 'Nellore'
 'patna' 'Meerut' 'hyderabad ' 'Sahibabad' 'Howrah' 'BHUBANESWAR'
 'Trichur' 'Ambala' 'Khopoli' 'keral' 'Roorkee' 'Greater NOIDA'
 'Navi mumbai' 'ghaziabad' 'Allahabad' 'Delhi/NCR' 'Panchkula '
```

```
'Ranchi '
 'Jalandhar' 'manesar' 'vapi' 'PILANI' 'muzzafarpur' 'RAS AL KHAIMAH'
 'bihar' 'singaruli' 'KANPUR' 'Banglore ' 'pondy' 'Mohali' 'Phagwara'
 ' Mumbai' ' bangalore' 'GURAGAON' 'Baripada' 'MEERUT' 'Yamuna Nagar'
 'shahibabad' 'sampla' 'Guwahati' 'Rourkela' 'Banaglore' 'Vellore'
'Dausa'
 'latur (Maharashtra )' 'NEW DELHI' 'kanpur' 'Mainpuri' 'karnal'
'Dammam'
 'Haldia' 'sambalpur' 'RAE BARELI' 'ranchi' 'jAipur' 'BANGLORE'
'Patiala'
 'Gorakhpur' 'new dehli' 'BANGALORE ' 'Ambala City' 'Karad' 'Rajpura'
 'Pilani' 'haryana' 'Asifabadbanglore']
******************** Gender ********************
['f' 'm']
******************** DOB ********************
['1990-02-19T00:00:00.000000000' '1989-10-04T00:00:00.000000000'
 '1992-08-03T00:00:00.000000000' ... '1986-02-28T00:00:00.000000000'
 '1990-06-22T00:00:00.000000000' '1987-04-15T00:00:00.000000000']
******************** 10percentage ********************
[84.3  85.4  85.   85.6  78.   89.92 86.08 92.   90.   77.   88.6  81.
 83.6  60.8  81.2  69.54 85.8  65.   79.   91.2  75.67 92.5  70.   88.
 86.8  90.88 76.8  84.   64.   77.2  87.   73.   71.   91.5  52.33
66.6
 75.   91.4  59.   84.5  95.8  69.6  86.6  81.92 66.5  80.   85.2
89.25
 58.4  90.8  89.88 63.3  72.   83.   85.3  60.66 89.37 84.4  89.   86.
 70.6  93.44 76.   86.4  84.83 72.3  69.   73.33 86.16 56.   79.25
88.66
 80.8  81.16 83.5  55.6  84.6  62.   68.4  51.   79.6  63.4  86.53
76.18
 91.43 76.17 82.4  58.   90.04 60.4  74.23 59.8  82.3  84.2  60.
82.8
 82.6  74.   89.8  88.67 64.66 77.8  89.5  64.56 91.12 73.08 78.33
81.33
 89.7  87.4  92.66 76.87 91.3  77.9  76.2  77.78 65.6  65.8  67.75
91.6
 78.66 78.4  61.   90.4  58.6  82.2  82.   90.1  86.17 83.89 76.7
88.2
 80.6  91.   74.4  79.28 84.96 92.8  79.4  66.8  79.8  65.3  94.6
83.33
 80.83 67.23 86.2  55.   86.62 69.13 89.12 87.5  66.4  80.33 75.2
50.6
 81.1  60.14 78.83 75.8  77.66 64.48 67.   75.6  85.28 71.5  93.6
93.33
 78.8  82.16 77.65 56.5  79.83 57.   89.42 72.8  86.3  77.5  71.2
80.2
 73.6  68.   74.7  69.2  65.33 96.48 82.5  91.8  93.4  68.5  73.4
72.2
 71.8  66.33 71.67 70.2  90.27 67.67 53.4  76.16 65.71 79.5  87.2
80.4
```

```
 88.4  90.6  89.2  81.6  86.5  77.16 72.33 79.89 75.4  72.83 88.33
78.88
 95.2  89.33 75.3  69.8  52.   58.2  90.5  70.83 62.13 74.5  63.
73.37
 83.4  92.6  78.6  76.4  86.88 66.   70.67 65.1  68.14 92.2  93.5
82.83
 81.66 90.15 44.16 94.4  64.4  85.17 70.1  88.25 94.8  77.4  66.66
81.03
 44.   45.6  87.8  72.6  79.86 84.67 48.   53.3  71.66 68.8  78.15
83.43
 86.9  84.8  75.06 86.66 70.9  81.12 67.5  78.2  71.06 71.28 62.1
90.56
 87.6  71.3  88.8  86.09 67.72 76.6  68.6  63.43 70.4  67.6  73.8
55.5
 74.67 92.16 83.66 62.86 49.   87.11 77.53 88.5  61.9  79.2  83.8
79.33
 67.3  83.2  72.5  77.67 94.2  59.33 87.63 89.76 84.14 88.17 59.6
64.3
 75.04 82.9  73.5  73.67 77.7  87.69 83.16 71.32 61.6  61.8  85.06
91.71
 75.46 74.8  67.4  90.3  86.15 64.7  69.7  82.33 58.83 75.36 76.5
66.67
 67.2  80.88 88.88 71.4  65.2  96.   61.2  81.4  68.66 65.56 86.7
63.2
 62.2  68.2  77.85 83.3  56.78 83.04 71.1  76.53 74.83 59.3  73.2
86.33
 64.83 72.1  61.1  86.83 70.8  94.   80.3  64.6  84.1  84.93 92.83
89.6
 80.5  83.82 77.57 51.2  66.46 82.67 61.4  69.4  90.24 75.75 90.83
89.1
 87.33 83.1  88.34 91.67 87.3  48.8  75.38 55.52 71.05 77.63 92.47 93.
 68.33 79.16 85.33 80.42 78.25 95.52 87.86 89.4  94.7  97.12 93.94
59.57
 80.53 81.9  63.6  66.16 62.5  69.5  80.93 82.7  79.29 81.5  62.34
83.68
 70.66 60.1  77.38 68.83 94.43 85.53 88.09 74.2  85.73 72.4  67.7
79.78
 81.3  79.37 85.5  87.83 70.33 80.64 58.7  60.2  77.81 85.67 89.9
79.75
 75.73 87.68 60.5  81.38 75.52 48.5  88.3  82.1  85.18 81.8  90.46
70.5
 79.52 89.29 61.75 78.67 71.6  80.13 81.67 95.   81.14 72.16 87.04
88.64
 82.56 90.01 63.66 65.17 91.84 92.1  43.   65.23 92.48 82.88 73.07
58.56
 85.83 67.34 85.75 80.7  79.23 70.26 52.7  75.86 90.16 90.2  78.5
58.9
 80.32 67.16 73.06 92.93 85.76 87.52 88.36 81.73 60.7  87.7  79.85
68.9
 73.83 64.8  77.84 74.14 64.28 92.4  73.94 63.33 70.06 88.7  89.3
```

```
83.83
 91.33 80.07 72.17 63.5  69.17 67.42 84.16 76.64 78.13 61.69 81.7
82.46
 64.57 78.1  66.3  59.71 80.67 77.88 88.53 93.38 52.93 78.17 72.36
84.75
 86.67 77.6  74.3  62.4  65.16 59.2  69.85 79.68 65.4  94.72 71.86
81.25
 87.62 54.   85.92 74.33 82.28 75.1  69.73 92.12 70.3  76.33 80.1
76.67
 77.83 57.67 83.14 50.   91.21 81.83 78.16 80.14 93.3  61.63 73.73
80.46
 76.48 82.25 56.16 57.88 87.07 67.1  87.73 77.12 64.53 86.46 82.62
53.06
 74.62 76.66 45.33 69.69 53.   75.33 74.28 74.6  74.88 74.53 46.24
80.15
 86.13 85.72 51.36 78.53 53.8  86.1  84.43 76.36 77.86 88.83 88.18
79.14
 54.5  84.33 78.3  77.44 87.31 58.16 61.3  77.33 75.12 72.12 65.26
83.56
 50.5  82.27 67.36 87.16 75.83 78.44 91.9  68.3  92.64 58.17 83.46
88.04
 70.58 71.17 84.13 64.62 66.15 67.8  57.78 56.2  61.5  94.1  74.16
78.93
 70.56 85.16 82.26 71.13 65.66 71.71 86.37 88.57 59.05 79.66 89.39
95.54
 55.3  70.14 87.23 85.38 86.92 95.04 95.6  60.83 71.33 94.16 78.57
80.16
 70.25 82.13 88.76 51.6  70.76 57.5  68.67 74.18 68.25 71.04 64.43
82.24
 90.06 67.12 75.85 87.81 76.93 65.5  92.3  50.66 61.83 63.16 64.2
53.6
 83.69 80.04 78.86 70.61 62.8  67.06 65.85 94.24 63.8  75.77 94.5
71.83
 91.1  91.52 90.08 93.16 86.34 88.1  97.   62.93 76.96 85.46 78.08
66.7
 95.86 92.09 81.46 81.86 61.73 77.22 46.8  79.3  72.75 93.8  93.67
75.16
 72.45 85.62 86.85 93.28 58.33 64.33 75.62 54.8  58.5  91.68 69.3
62.6
 65.41 91.62 88.16 92.32 69.83 83.67 69.92 52.5  56.4  54.6  78.61
69.53
 75.71 71.84 70.16 69.66 95.4  84.62 91.53 85.66 61.57 80.63 69.33
62.3
 87.87 70.75 87.58 58.8  62.88 73.66 97.76 88.44 54.83 67.25 90.76
82.75
 75.66 91.04 90.58 91.86 73.1  73.3  69.1  51.83 62.26 65.67 68.65
51.42
 78.28 80.58 55.33 91.17 74.9  93.43 90.81 89.89 62.67 90.26 62.15
70.08
 87.88 72.67 93.2  60.46 71.07 46.   54.4  86.12 72.15 71.85 49.9
```

```
83.75
 71.73 90.33 86.14 66.2  88.75 67.33 57.14 67.76 82.66 77.92 79.38
67.17
 89.17 79.67 96.8  71.37 82.87 89.44 71.76 57.7  89.23 79.06 83.25
61.86
 89.56 82.14 70.27 59.7  93.07 79.9  64.5  66.85 69.16 52.09 78.72]
******************** 10board ********************
['board ofsecondary education,ap' 'cbse' 'state board' 'mp board
bhopal'
 'icse' 'karnataka secondary school of examination' 'up'
 'karnataka state education examination board' 'ssc'
 'kerala state technical education' 0 'bseb'
 'state board of secondary education, andhra pradesh' 'matriculation'
 'gujarat state board' 'karnataka state board' 'wbbse'
 'maharashtra state board' 'icse board' 'up board'
 'board of secondary education(bse) orissa'
 'little jacky matric higher secondary school' 'uttar pradesh board'
 'bsc,orissa' 'mp board' 'upboard' 'matriculation board' 'j & k bord'
 'rbse' 'central board of secondary education' 'pseb' 'jkbose'
 'haryana board of school education,(hbse)' 'metric' 'ms board'
'kseeb'
 'stateboard' 'maticulation' 'karnataka secondory education board'
 'mumbai board' 'sslc' 'kseb' 'board secondary  education' 'matric
board'
 'board of secondary education' 'west bengal board of secondary
education'
 'jharkhand secondary examination board,ranchi' 'u p board'
'bseb,patna'
 'hsc' 'bse' 'sss pune' 'karnataka education board (keeb)' 'kerala'
 'state board of secondary education( ssc)' 'gsheb' 'up(allahabad)'
 'nagpur' 'don bosco maatriculation school'
 'karnataka state secondary education board' 'maharashtra'
 'karnataka secondary education board'
 'himachal pradesh board of school education'
 'certificate of middle years program of ib'
 'karnataka board of secondary education'
 'board of secondary education rajasthan' 'uttarakhand board' 'ua'
 'board of secendary education orissa'
 'karantaka secondary education and examination borad' 'hbsc'
 'kseeb(karnataka secondary education examination board)'
 'cbse[gulf zone]' 'hbse' 'state(karnataka board)'
 'jharkhand accademic council'
 'jharkhand secondary examination board (ranchi)'
 'karnataka secondary education examination board' 'delhi board'
 'mirza ahmed ali baig' 'jseb' 'bse, odisha' 'bihar board'
 'maharashtra state(latur board)' 'rajasthan board' 'mpboard'
'upbhsie'
 'secondary board of rajasthan' 'tamilnadu matriculation board'
 'jharkhand secondary board'
```

```
 'board of secondary education,andhara pradesh' 'up baord' 'state'
 'board of intermediate education'
 'state board of secondary education,andhra pradesh'
 'up board , allahabad' 'stjosephs girls higher sec school,dindigul'
 'maharashtra board' 'education board of kerala' 'board of ssc'
 'maharashtra state board pune' 'board of school education harayana'
 'secondary school cerficate' 'maharashtra sate board' 'ksseb'
 'bihar examination board, patna' 'latur'
 'board of secondary education, rajasthan' 'state borad hp' 'cluny'
 'bsepatna' 'up borad' 'ssc board of andrapradesh' 'matric'
'bse,orissa'
 'ssc-andhra pradesh' 'mp' 'karnataka education board' 'mhsbse'
 'karnataka sslc board bangalore' 'karnataka' 'u p'
 'secondary school of education' 'state board of karnataka'
 'karnataka secondary board' 'andhra pradesh board ssc'
 'stjoseph of cluny matrhrsecschool,neyveli,cuddalore district'
 'hse,orissa' 'national public school' 'nagpur board'
 'jharkhand academic council' 'bsemp'
 'board of secondary education, andhra pradesh'
 'board of secondary education orissa'
 'board of secondary education,rajasthan(rbse)'
 'board of secondary education,ap'
 'board of secondary education,andhra pradesh'
 'jawahar navodaya vidyalaya' 'aisse'
 'karnataka board of higher education' 'bihar' 'kerala state board'
 'cicse' 'tn state board' 'kolhapur divisional board, maharashtra'
 'bharathi matriculation school' 'uttaranchal state board' 'wbbsce'
 'mp state board' 'seba(assam)' 'anglo indian' 'gseb' 'uttar pradesh'
 'ghseb' 'board of school education uttarakhand' 'msbshse,pune'
 'tamilnadu state board' 'kerala university'
 'uttaranchal shiksha avam pariksha parishad'
 'bse(board of secondary education)' 'bright way college, (up board)'
 'school secondary education, andhra pradesh'
 'secondary state certificate'
 'maharashtra state board of secondary and higher secondary
education,pune'
 'andhra pradesh state board' 'stmary higher secondary' 'cgbse'
 'secondary school certificate' 'rajasthan board ajmer' 'mpbse'
 'pune board' 'cbse ' 'board of secondary education,orissa'
 'maharashtra state board,pune' 'up bord'
 'kiran english medium high school' 'state board (jac, ranchi)'
 'gujarat board' 'state board ' 'sarada high scchool'
 'kalaimagal matriculation higher secondary school' 'karnataka board'
 'maharastra board' 'sslc board' 'ssc maharashtra board'
 'tamil nadu state' 'uttrakhand board'
 'bihar secondary education board,patna'
 'haryana board of school education'
 'sri kannika parameswari highier secondary school, udumalpet'
 'ksseb(karnataka state board)' 'nashik board'
```

```
 'jharkhand secondary education board' 'himachal pradesh board'
 'maharashtra satate board'
 'maharashtra state board mumbai divisional board'
 'dav public school,hehal' 'state board of secondary education, ap'
 'rajasthan board of secondary education' 'hsce'
 'karnataka secondary education' 'board of secondary education,odisha'
 'maharashtra nasik board'
 'west bengal board of secondary examination (wbbse)'
 'holy cross matriculation hr sec school' 'cbsc' 'apssc' 'bseb patna'
 'kolhapur' 'bseb, patna' 'up board allahabad' 'biharboard'
 'nagpur board,nagpur' 'pune' 'gyan bharati school' 'rbse,ajmer'
 'board of secondaray education' 'secondary school education' 'state
bord'
 'jbse,jharkhand' 'hse' 'madhya pradesh board'
 'bihar school examination board'
 'west bengal board of secondary eucation' 'state boardmp board '
 'icse board , new delhi' 'board of secondary education (bse) orissa'
 'maharashtra state board for ssc' 'board of secondary school
education'
 'latur board' "stmary's convent inter college" 'nagpur divisional
board'
 'ap state board' 'cgbse raipur' 'uttranchal board' 'ksbe'
 'central board of secondary education, new delhi'
 'bihar school examination board patna' 'cbse board' 'sslc,karnataka'
 'mp-bse' 'up bourd' 'dav public school sec 14'
 'board of school education haryana'
 'council for indian school certificate examination' 'aurangabad
board'
 'j&k state board of school education'
 'maharashtra state board of secondary and higher secondary education'
 'maharashtra state boar of secondary and higher secondary education'
 'ssc regular' 'karnataka state examination board' 'nasik'
 'west bengal  board of secondary education' 'up board,allahabad'
 'bseb ,patna'
 'state board - west bengal board of secondary education : wbbse'
 'maharashtra state board of secondary & higher secondary education'
 'delhi public school' 'karnataka secondary eduction'
 'secondary education board of rajasthan' 'maharashtra board, pune'
 'rbse (state board)' 'apsche' 'board of  secondary education'
 'board of high school and intermediate education uttarpradesh' 'kea'
 'board of secondary education - andhra pradesh'
 'ap state board for secondary education' 'seba'
 'punjab school education board, mohali' 'jharkhand acedemic council'
 'hse,board' 'board of ssc education andhra pradesh' 'up-board'
 'bse,odisha']
********** 12graduation **********
[2007 2010 2008 2009 2006 2011 2005 1995 2004 2012 2003 2002 2001 1998
 2013 1999]
********** 12percentage **********
```

```
[95.8  85.   68.2  83.6  76.8  87.   67.5  91.   91.2  72.2  83.7  86.
 69.83 62.4  79.9  64.43 74.8  66.66 64.8  62.2  84.63 74.4  95.2
43.42
 90.   82.8  82.5  83.   68.   74.   92.   86.1  84.4  68.4  61.
93.8
 85.4  67.   89.66 68.6  60.   73.   87.7  87.16 82.   71.4  75.
61.46
 49.5  72.5  81.5  78.   90.1  70.1  95.4  61.1  49.   79.   88.3
77.8
 94.4  86.67 73.2  64.   77.   78.67 72.   78.2  95.   82.4  60.2
62.6
 70.   71.33 70.08 56.   80.   84.33 87.9  65.   68.5  94.2  66.   88.
 76.6  50.8  66.5  76.4  65.8  48.   63.   71.55 86.33 71.3  57.6
83.4
 75.16 77.77 60.25 89.6  86.4  80.4  59.   73.6  63.6  66.6  86.8
79.6
 87.58 81.4  89.   62.   47.   77.2  71.2  54.   67.6  91.9  63.1  69.
 68.46 85.2  78.4  82.2  95.6  81.33 88.9  82.75 72.6  66.67 70.2
61.5
 70.6  79.4  61.8  95.5  80.7  60.4  77.14 75.2  81.2  80.8  88.88
83.9
 65.2  83.1  80.6  70.16 90.91 84.7  68.55 59.16 78.83 59.9  75.6
87.8
 79.2  80.3  82.1  65.5  84.3  64.4  91.6  95.3  69.8  86.9  73.4
56.9
 86.7  64.7  80.5  61.6  90.7  81.   57.   92.6  78.6  71.   71.5
70.4
 89.5  76.66 80.1  54.4  80.9  84.8  69.45 93.4  56.8  91.5  90.67
64.5
 96.1  62.8  94.1  89.7  76.   73.8  78.3  96.6  60.17 75.4  72.4  52.
 77.4  69.6  56.2  78.43 82.7  74.45 76.2  68.8  78.8  50.   67.4
95.41
 84.   71.08 94.5  67.75 87.1  81.8  80.2  83.58 62.83 69.4  86.5
67.67
 74.2  66.4  55.   61.2  64.6  43.12 61.57 84.5  83.8  89.1  74.6
81.1
 80.83 67.8  76.5  87.83 69.9  88.4  58.55 83.33 82.6  69.2  63.2
88.7
 62.66 61.7  96.7  62.5  79.57 93.   78.16 74.92 94.16 51.4  97.1
94.9
 87.4  63.4  87.66 65.4  68.67 58.   68.66 67.7  91.3  79.8  81.6
67.25
 55.2  65.6  64.3  51.63 93.7  69.04 50.2  96.5  77.7  90.2  72.67
73.83
 71.9  82.83 82.08 77.38 71.8  66.8  71.25 53.   76.62 72.77 79.83
77.3
 81.17 81.12 62.16 67.68 57.8  85.3  83.67 75.69 75.9  55.02 77.5
88.83
 75.5  92.3  67.2  62.3  72.48 59.2  72.8  60.13 94.3  89.8  81.3
```

82.02
 76.89 60.16 77.6 69.33 87.5 74.5 46.   74.88 68.33 84.67 87.2
66.3
 58.6 85.36 68.7 55.54 68.89 82.3 58.2 48.34 54.2 92.7 71.01
53.8
 56.6 66.77 71.6 90.6 82.66 89.08 93.6 72.3 69.7 84.17 92.1
90.4
 92.67 64.33 62.81 67.17 95.1 68.83 97.8 88.1 48.8 60.05 69.07
89.3
 86.3 94.6 86.25 79.23 97.   84.75 89.2 77.54 60.8 84.6 86.2
65.16
 94.33 74.67 94.7 92.2 90.33 73.61 53.6 58.4 78.86 76.44 78.66
97.5
 62.26 69.84 78.13 60.83 51.3 57.67 67.9 56.3 88.6 52.5 76.33 94.
 64.88 80.75 71.66 50.3 76.67 67.57 64.45 69.5 72.15 59.77 61.4
72.1
 90.25 97.4 89.4 57.5 70.8 56.4 96.2 88.77 96.   83.2 69.17
80.33
 79.04 59.8 81.66 72.46 92.5 50.5 96.8 93.5 65.33 92.4 72.66
87.33
 94.8 60.3 90.3 61.17 82.25 88.5 75.11 70.33 85.8 88.8 45.6
58.33
 76.24 45.   68.3 89.75 69.67 88.2 65.56 74.12 86.54 57.11 65.66
85.67
 96.3 72.31 70.3 77.86 94.75 69.32 84.9 60.1 86.6 70.15 66.62
79.16
 76.7 65.25 93.3 68.15 80.25 76.77 65.9 91.25 54.14 55.55 95.65
75.91
 66.2 70.66 59.4 58.92 51.   63.7 85.5 70.83 81.7 63.8 75.25 43.
 61.16 56.12 60.66 69.58 70.04 79.67 49.67 76.3 75.67 61.3 82.33
88.33
 88.45 96.75 84.25 78.5 77.85 82.46 74.3 73.67 60.01 91.33 69.16
91.7
 65.83 74.83 58.5 72.17 87.6 75.85 79.87 85.9 85.6 85.7 73.11
88.66
 82.56 78.34 89.33 66.15 87.25 69.53 53.17 79.19 77.16 54.08 52.2
83.5
 85.57 60.33 67.23 47.6 73.25 69.88 84.53 89.9 79.3 93.1 92.25
91.58
 75.14 73.3 77.1 79.5 59.6 50.4 74.89 63.33 77.23 83.43 83.3
95.7
 90.8 75.75 47.2 86.31 85.17 63.83 64.57 73.69 66.86 81.75 55.6
62.12
 80.58 74.96 70.14 69.54 61.83 72.33 74.33 89.04 71.67 60.36 90.66
60.5
 70.26 57.2 87.69 85.56 64.2 45.5 70.5 93.9 57.58 85.1 64.1
79.1
 91.83 88.93 46.33 94.17 95.9 81.9 76.12 84.28 55.4 58.04 91.8
74.7

```
 83.25 91.4  54.66 82.9  66.46 71.52 56.17 59.66 62.7  78.25 61.92
68.16
 53.83 47.83 92.75 63.44 87.05 92.08 68.69 55.33 89.67 63.77 74.53
60.6
 77.83 56.04 70.67 65.1  87.3  56.22 96.4  83.62 74.44 97.9  57.33
71.17
 95.08 84.2  97.6  57.83 93.41 79.33 64.83 75.66 75.8  62.33 92.9
92.8
 70.92 81.42 90.9  52.7  82.03 73.63 75.1  87.17 73.33 62.23 61.33
52.9
 83.88 83.75 83.83 67.74 59.38 68.04 77.81 89.26 74.71 53.2  53.85
65.54
 51.6  68.77 86.29 88.75 49.6  77.33 58.8  71.85 66.17 53.44 66.83
93.2
 85.33 72.9  57.1  98.7  88.16 54.33 85.42 90.5  74.16 98.2  82.71
90.03
 85.23 63.16 96.33 84.09 74.05 81.16 40.   60.06 89.16 69.89 61.23
79.82
 65.72 55.8  76.85 69.3  90.17 53.16 61.85 89.58 83.23 67.3  93.33
52.8
 73.17 52.34 63.5  77.11 62.11 97.2  78.88 65.12 55.44 53.55 80.22
62.9
 55.66 94.91 76.83 64.08 83.34 85.88 63.3  53.4  54.8  54.5  74.25
79.75
 60.44 85.66 55.5  51.23 90.83 67.1  83.16 79.7  56.1  68.17 82.53
87.14
 63.9  74.14 89.91 77.06 60.42 75.33 96.25 69.12 77.56 86.46 81.26
64.31
 71.83 86.91 81.25 54.83 59.11 91.1  81.67 53.33 82.55]
********** 12board **********
['board of intermediate education,ap' 'cbse' 'state board' 'mp board'
 'isc' 'icse' 'karnataka pre university board' 'up' 'p u board,
karnataka'
 'dept of pre-university education' 'bie' 'kerala state hse board'
 'up board' 0 'bseb' 'chse' 'puc' ' upboard'
 'state  board of intermediate education, andhra pradesh'
 'karnataka state board'
 'west bengal state council of technical education' 'wbchse'
 'maharashtra state board' 'ssc' 'isc board'
 'sda matric higher secondary school' 'uttar pradesh board' 'ibe'
'chsc'
 'board of intermediate' 'isce' 'upboard' 'sbtet'
 'hisher seconadry examination(state board)' 'pre university'
 'borad of intermediate' 'j & k board'
 'intermediate board of andhra pardesh' 'rbse'
 'central board of secondary education' 'jkbose' 'hbse'
 'board of intermediate education' 'state' 'ms board' 'pue'
 'intermediate state board' 'stateboard' 'hsc'
 'electonincs and communication(dote)' 'karnataka pu board'
```

```
 'government polytechnic mumbai , mumbai board' 'pu board'
 'baord of intermediate education' 'apbie' 'andhra board'
 'tamilnadu stateboard'
 'west bengal council of higher secondary education' 'cbse,new delhi'
 'u p board' 'intermediate' 'biec,patna'
 'diploma in engg (e &tc) tilak maharashtra vidayapeeth' 'hsc pune'
 'pu board karnataka' 'kerala' 'gsheb' 'up(allahabad)' 'nagpur'
 'st joseph hr sec school' 'pre university board' 'ipe' 'maharashtra'
 'kea' 'apsb' 'himachal pradesh board of school education' 'staae
board'
 'international baccalaureate (ib) diploma' 'nios'
 'karnataka board of university' 'board of secondary education
rajasthan'
 'uttarakhand board' 'ua' 'scte vt orissa' 'matriculation'
 'department of pre-university education' 'wbscte'
 'preuniversity board(karnataka)' 'jharkhand accademic council'
'bieap'
 'msbte (diploma in computer technology)'
 'jharkhand acamedic council (ranchi)'
 'department of pre-university eduction' 'biec' 'all india board'
'sjrcw'
 ' board of intermediate' 'msbte' 'sri sankara vidyalaya' 'chse,
odisha'
 'bihar board' 'maharashtra state(latur board)' 'rajasthan board'
 'mpboard' 'state board of technical eduction panchkula' 'upbhsie'
'apbsc'
 'state board of technical education and training'
 'secondary board of rajasthan'
 'tamilnadu higher secondary education board' 'jharkhand academic
council'
 'board of intermediate education,hyderabad' 'up baord' 'pu' 'dte'
 'board of secondary education' 'pre-university'
 'board of intermediate education,andhra pradesh' 'up board ,
allahabad'
 'srv girls higher sec school,rasipuram'
 'intermediate board of education,andhra pradesh'
 'intermediate board examination'
 'department of pre-university education, bangalore'
 'stmiras college for girls' 'mbose'
 'department of pre-university education(government of karnataka)'
'dpue'
 'msbte pune' 'board of school education harayana' 'sbte, jharkhand'
 'bihar intermediate education council, patna' 'higher secondary'
 's j polytechnic' 'latur' 'board of secondary education, rajasthan'
 'jyoti nivas' 'pseb' 'biec-patna'
 'board of intermediate education,andra pradesh' 'chse,orissa'
 'pre-university board' 'mp' 'intermediate board'
 'govt of karnataka department of pre-university education'
 'karnataka education board' 'board of secondary school of education'
```

```
 'pu board ,karnataka' 'karnataka secondary education board'
 'karnataka sslc' 'board of intermediate ap' 'u p'
 'state board of karnataka' 'directorate of technical
education,banglore'
 'matric board' 'andhpradesh board of intermediate education'
 'stjoseph of cluny matrhrsecschool,neyveli,cuddalore district' 'bte
up'
 'scte and vt ,orissa' 'hbsc' 'jawahar higher secondary school'
 'nagpur board' 'bsemp' 'board of intermediate education, andhra
pradesh'
 'board of higher secondary orissa'
 'board of secondary education,rajasthan(rbse)'
 'board of intermediate education:ap,hyderabad' 'science college'
 'karnatak pu board' 'aissce' 'pre university board of karnataka'
'bihar'
 'kerala state board' 'uo board' 'cicse' 'karnataka board'
 'tn state board' 'kolhapur divisional board, maharashtra'
 'jaycee matriculation school'
 'board of higher secondary examination, kerala' 'uttaranchal state
board'
 'intermidiate' 'bciec,patna' 'bice' 'karnataka state' 'state broad'
 'wbbhse' 'gseb' 'uttar pradesh' 'ghseb'
 'board of school education uttarakhand' 'gseb/technical education
board'
 'msbshse,pune' 'tamilnadu state board' 'board of technical education'
 'kerala university' 'uttaranchal shiksha avam pariksha parishad'
 'chse(concil of higher secondary education)'
 'bright way college, (up board)' 'board of intermidiate'
 'higher secondary state certificate' 'karanataka secondary board'
 'maharashtra board' 'andhra pradesh state board' 'cgbse'
 'diploma in computers' 'bte,delhi' 'rajasthan board ajmer' 'mpbse'
 'pune board' 'state board of technical education' 'gshseb'
 'amravati divisional board' 'dote (diploma - computer engg)' 'up
bord'
 'karnataka pre-university board' 'jharkhand board'
 'punjab state board of technical education & industrial training'
 'department of technical education' 'sri chaitanya junior kalasala'
 'state board (jac, ranchi)' 'gujarat board' 'aligarh muslim
university'
 'tamil nadu state board' 'hse' 'karnataka secondary education'
 'state board ' 'karnataka pre unversity board'
 'ks rangasamy institute of technology'
 'karnataka board secondary education' 'narayana junior college'
'bteup'
 'board of intermediate(bie)' 'hsc maharashtra board' 'tamil nadu
state'
 'uttrakhand board' 'psbte' 'stateboard/tamil nadu'
 'intermediate council patna' 'technical board, punchkula'
 'board of intermidiate examination'
```

```
 'sri kannika parameswari highier secondary school, udumalpet' 'ap
board'
 'nashik board' 'himachal pradesh board' 'maharashtra satate board'
 'andhra pradesh board of secondary education' 'tamil nadu
polytechnic'
 'maharashtra state board mumbai divisional board'
 'department of pre university education' 'dav public school,hehal'
 'board of intermediate education, ap'
 'rajasthan board of secondary education'
 'department of technical education, bangalore' 'chse,odisha'
 'maharashtra nasik board'
 'west bengal council of higher secondary examination (wbchse)'
 'holy cross matriculation hr sec school' 'cbsc' 'pu  board karnataka'
 'biec patna' 'kolhapur' 'bseb, patna' 'up board allahabad'
'intermideate'
 'nagpur board,nagpur' 'diploma(msbte)' 'dav public school'
 'pre university board, karnataka' 'ssm srsecschool' 'state bord'
 'jstb,jharkhand' 'intermediate board of education' 'mp board bhopal'
 'pub' 'madhya pradesh board' 'bihar intermediate education council'
 'west bengal council of higher secondary eucation'
 'isc board , new delhi' 'mpc'
 'certificate for higher secondary education (chse)orissa'
 'maharashtra state board for hsc' 'board of intermeadiate education'
 'latur board' 'andhra pradesh' 'karnataka pre-university'
 'lucknow public college' 'nagpur divisional board'
 'ap intermediate board' 'cgbse raipur' 'uttranchal board' 'jiec'
 'central board of secondary education, new delhi'
 'bihar school examination board patna'
 'state board of technical education harayana' 'mp-bse' 'up bourd'
 'dav public school sec 14'
 'haryana state board of technical education chandigarh'
 'council for indian school certificate examination'
 'jaswant modern school' 'madhya pradesh open school' 'aurangabad
board'
 'j&k state board of school education'
 'diploma ( maharashtra state board of technical education)'
 'board of technicaleducation ,delhi'
 'maharashtra state boar of secondary and higher secondary education'
 'hslc (tamil nadu state board)' 'karnataka state examination board'
 'puboard' 'nasik' 'west bengal board of higher secondary education'
 'up board,allahabad' 'board of intrmediate education,ap' 'cbese'
 'karnataka state pre- university board'
 'state board - west bengal council of higher secondary education :
wbchse'
 'maharashtra state board of secondary & higher secondary education'
 'biec, patna' 'state syllabus' 'cbse board' 'scte&vt'
 'board of intermediate,ap'
 'secnior secondary education board of rajasthan'
 'maharashtra board, pune' 'rbse (state board)'
```

```
 'board of intermidiate education,ap'
 'board of high school and intermediate education uttarpradesh'
 'higher secondary education' 'board fo intermediate education, ap'
 'intermedite' 'ap board for intermediate education' 'ahsec'
 'punjab state board of technical education & industrial training,
chandigarh'
 'state board - tamilnadu' 'jharkhand acedemic council'
 'scte & vt (diploma)' 'karnataka pu' 'board of intmediate education
ap'
 'up-board' 'boardofintermediate']
********** CollegeID **********
[1141 5807   64 ... 3572 6327 4883]
********** CollegeTier **********
[2 1]
********** Degree **********
['B.Tech/B.E.' 'MCA' 'M.Tech./M.E.' 'M.Sc. (Tech.)']
********** Specialization **********
['computer engineering' 'electronics and communication engineering'
 'information technology' 'computer science & engineering'
 'mechanical engineering' 'electronics and electrical engineering'
 'electronics & telecommunications'
 'instrumentation and control engineering' 'computer application'
 'electronics and computer engineering' 'electrical engineering'
 'applied electronics and instrumentation'
 'electronics & instrumentation eng' 'information science engineering'
 'civil engineering' 'mechanical and automation'
 'industrial & production engineering'
 'control and instrumentation engineering' 'metallurgical engineering'
 'electronics and instrumentation engineering' 'electronics
engineering'
 'ceramic engineering' 'chemical engineering' 'aeronautical
engineering'
 'other' 'biotechnology' 'embedded systems technology'
 'electrical and power engineering' 'computer science and technology'
 'mechatronics' 'automobile/automotive engineering' 'polymer
technology'
 'mechanical & production engineering' 'power systems and automation'
 'instrumentation engineering' 'telecommunication engineering'
 'industrial & management engineering' 'industrial engineering'
 'computer and communication engineering'
 'information & communication technology' 'information science'
 'internal combustion engine' 'computer networking'
 'biomedical engineering' 'electronics' 'computer science']
********** collegeGPA **********
[78.   70.06 70.   ... 65.05 74.73 70.42]
********** CollegeCityID **********
[1141 5807   64 ... 3572 6327 4883]
********** CollegeCityTier **********
[0 1]
```

```
******************** CollegeState ********************
['Andhra Pradesh' 'Madhya Pradesh' 'Uttar Pradesh' 'Delhi' 'Karnataka'
 'Tamil Nadu' 'West Bengal' 'Maharashtra' 'Haryana' 'Telangana'
'Orissa'
 'Punjab' 'Kerala' 'Gujarat' 'Rajasthan' 'Chhattisgarh' 'Uttarakhand'
 'Jammu and Kashmir' 'Jharkhand' 'Himachal Pradesh' 'Bihar' 'Assam'
'Goa'
 'Sikkim' 'Union Territory' 'Meghalaya']
******************** GraduationYear ********************
[2011 2012 2014 2016 2013 2010 2015 2009 2017    0 2007]
******************** English ********************
[515 695 615 635 545 560 590 605 565 495 380 395 485 685 465 455 385
370
 625 575 415 535 580 475 570 430 450 510 425 555 300 505 440 525 420
640
 444 630 665 675 325 405 375 315 710 345 250 350 275 360 265 595 585
520
 500 735 765 335 490 660 355 530 365 655 730 445 720 645 650 875 534
454
 544 295 285 435 464 705 554 745 280 825 290 715 310 215 700 870 305
524
 755 790 800 205 725 780 404 770 805 180 830 795 255 324 775 394 240
225
 850 684 334]
******************** Logical ********************
[585 610 545 625 555 435 670 565 455 605 580 425 520 530 495 445 535
360
 335 510 570 375 405 485 475 525 640 595 560 340 395 415 465 505 385
460
 410 500 645 480 355 450 440 470 255 305 590 630 365 350 325 400 205
655
 295 345 390 665 515 540 680 245 620 420 575 635 554 315 615 215 370
300
 274 685 324 675 650 464 684 275 334 544 454 534 404 795 285 715 700
674
 690 695 394 270 705 310 490 330 280 735 380 290 265 240 195 235 660]
******************** Quant ********************
[525 780 370 625 465 620 380 590 530 545 565 715 470 645 355 515 435
445
 485 270 630 575 405 605 385 695 450 295 430 415 635 475 460 825 500
455
 554 595 495 665 250 310 325 390 510 535 340 440 705 534 400 395 570
750
 330 320 454 365 615 505 425 235 210 585 810 555 735 560 524 690 870
765
 675 520 655 305 725 840 650 375 720 265 280 464 404 800 680 260 674
760
 345 335 165 685 544 215 180 795 200 860 334 285 514 195 494 214 275
315
 324 175 684 225 740 805 444 410 135 255 220 755 855 145 245 885 120
```

```
900
 794 775 745 504 820 150 710 190 185 155 580 394]
******************** Domain ********************
[ 0.63597876  0.96060325  0.45087658  0.97439611  0.12450207 -1.
  0.35653649  0.8295846   0.69447933  0.49359639  0.76567358
0.9682375
  0.22948175  0.53838689  0.30840058  0.91139528  0.56326782
0.86468541
  0.64938971  0.74475835  0.88412251  0.88162007  0.20739217
0.48674701
  0.67074315  0.62264292  0.41383826  0.52592258  0.73579571
0.13044174
  0.23780284  0.11213944  0.37755142  0.06696071  0.08005528
0.92564577
  0.84312373  0.91686996  0.78330354  0.60005718  0.79293628
0.79358061
  0.16563309  0.75537512  0.99990456  0.33878635  0.91077016
0.98205712
  0.84224832  0.01854094  0.05316031  0.94211655  0.12301673
0.48834798
  0.37605959  0.0587928   0.10487136  0.60064396  0.70409041
0.14478989
  0.81941653  0.65576694  0.02106623  0.44461772  0.83762073
0.72598415
  0.95389978  0.04099931  0.02196911  0.3423149   0.53586282
0.90148957
  0.96177212  0.67964464  0.93839914  0.19015341  0.99546472
0.99000876
  0.97879929  0.41433743  0.8246664   0.49063696  0.96600692
0.43696265
  0.99225919  0.97629256  0.98546139  0.90915194  0.86372418
0.07454627
  0.66183448  0.93037061  0.95224557  0.96090309  0.97706647
0.55738951
  0.7908818   0.99966434  0.8799152   0.85882669  0.97166349
0.19376844
  0.24545566  0.98374997  0.98720709  0.97952174  0.86873659
0.45901584
  0.02231329  0.94513486  0.27604723  0.53678267  0.4239514
0.29876913
  0.20671062  0.12569018  0.16363093  0.99138693  0.94327216
0.60703404
  0.99139758  0.25577819  0.96221701  0.99123063  0.87550391
0.02331217
  0.44714789  0.14325654  0.84084097  0.33618507  0.30451091
0.27925851
  0.84980271  0.03293702  0.65410747  0.28677781  0.21678549
0.95868202
  0.32874638  0.98466189  0.99674449  0.99434211  0.70082616
```

```
0.99887611
  0.86262547  0.60553355  0.68456507  0.71452936  0.44619825
0.92514031
  0.99526571  0.59828069  0.98652502  0.11166096  0.11655256
0.8957765
  0.1559085   0.99764277  0.88790332  0.88570369  0.99405087
0.02781507
  0.90394078  0.93174969  0.09575376  0.1925871   0.99868025
0.03114969
  0.61061199  0.75837988  0.63958738  0.27350033  0.0109953
0.21325143
  0.99925029  0.06844575  0.04222313  0.77657832  0.75618011
0.99808656
  0.99982917  0.01656464  0.00815478  0.89984384  0.92879365
0.99769826
  0.9956138   0.9788553   0.03056599  0.04219243  0.18477158
0.92460994
  0.81062051  0.99742783  0.66033499  0.99615597  0.99859145
0.00275015
  0.51986392  0.99729897  0.20273159  0.38009183  0.52173598
0.0139623
  0.24249992  0.29429542  0.4099522   0.80558321  0.24320866
0.96532716
  0.4845907   0.97524659  0.1798739   0.37113867  0.86613982
0.11502273
  0.99991041  0.99179243  0.22105882  0.0417332   0.06222129
0.83968608
  0.41552456  0.25302837  0.99657055  0.97289871  0.99896694
0.99391731
  0.52911566  0.55321638  0.01770484  0.91541843  0.2995973
0.66732645
  0.27845741  0.00853725  0.99858824  0.28278814  0.08874741
0.79984821
  0.99653553  0.90474069  0.93858826]
********** ComputerProgramming **********
[445  -1 395 615 645 405 735 385 485 605 495 355 515 545 425 525 455
475
 565 535 335 345 465 415 435 155 375 555 305 315 804 285 575 505 195
225
 595 275 334 365 685 655 625 585 665 325 235 255 205 494 695 635 215
464
 295 394 245 715 265 135 105 524 165 175 125 675 454 745 185 214 145
544
 725 840 404 755 705 115 554]
********** ElectronicsAndSemicon **********
[ -1 466 233 366 324 266 333 356 420 260 228 388 300 292 433 196 200
164
 400 484 500 452 516 166 533 566 612 133 548]
********** ComputerScience **********
```

```
[ -1 407 346 376 500 438 315 253 469 192 530 284 223 561 684 592 623
653
 130 715]
******************** MechanicalEngg ********************
[ -1 469 313 286 253 366 446 206 438 332 393 383 260 561 553 376 526
284
 409 473 340 223 420 538 346 435 512 407 580 280 358 500 315 254 616
564
 233 306 461 180 606 623]
******************** ElectricalEngg ********************
[ -1 484 606 393 500 553 580 446 420 324 388 356 313 633 516 366 612
452
 526 548 228 433 473 676 292 660 411 286 340 260 206]
******************** TelecomEngg ********************
[ -1 206 313 420 260 393 366 446 324 340 286 473 484 452 233 292 526
153
 516 356 548 228 196 164 388 500]
******************** CivilEngg ********************
[ -1 320 400 388 260 440 356 292 500 200 300 452 322 340 166 277 516
380
 433 280 420 460 480]
******************** conscientiousness ********************
[ 0.9737 -0.7335  0.2718  0.0464 -0.881  -0.3027  1.7081 -0.0154 -
0.159
 -1.308  -2.272   0.1282  0.3555  0.7027  1.7465  1.1336  0.8463
0.8192
 -0.1082 -1.0355 -0.4463  0.4155  0.99   -3.1994 -0.4173  1.5644 -
0.4854
 -1.0208  0.3941 -0.8772  0.51   -0.5899 -2.5039  1.2828  0.335  -
0.3014
  1.8517 -1.1644 -2.2351  0.6646 -0.2628 -1.8825 -1.4517  0.5591
1.4208
 -0.7264 -0.5116 -1.7389  0.2009 -0.0696 -2.5811 -2.3134  1.2772 -
2.8879
  1.4374 -1.3447  0.1623  1.7156 -1.9629 -2.457   1.9953 -2.0262 -
2.1175
 -2.7443 -1.4606  0.8578 -1.1901 -0.7651 -0.5719 -2.1698 -1.8083
1.592
 -0.9969 -1.3742 -1.4992 -1.5953 -3.6631  1.1283 -3.606   0.1788
0.2782
 -1.6538 -3.3539 -1.1128 -3.3188  0.4285  0.7419 -0.6491 -0.51   -
0.5236
 -0.9653 -3.1752  0.7208  1.3215 -1.6924 -0.1855 -0.6749  0.2318
1.5533
  1.0768  1.3686 -0.5332 -0.2632 -1.2287 -1.5765  1.9011  1.0896
0.215
 -3.4624  0.3836 -3.0448 -2.6007  1.2056 -1.295  -2.7357 -0.0415
1.6692
  0.626  -2.4266 -2.1561 -2.8903 -4.1267  0.4034 -1.9243 -1.3025 -
```

```
1.5964
  1.7852  0.6696  0.5522  0.8479 -1.0135 -0.1982 -3.8933 -3.5085 -
3.7496
 -4.0369 -1.977  -3.0315  1.2266 -0.4595  0.8986]
******************** agreeableness ********************
[ 0.8128  0.3789  1.7109  0.3448 -0.2793 -0.6201 -0.1054  1.2114
0.5454
  1.1248  0.0328  0.7119  1.9048  1.0449  0.2668  0.9688 -0.5913 -
2.1186
  0.8027  1.2028  0.1888  1.3779 -1.8393  0.6568 -0.4536 -0.5213
0.2124
  1.2808 -1.1196  0.2578 -2.4516  1.7488 -0.1206  0.0924 -0.0842 -
0.4353
  1.5444 -0.9531 -2.6847 -0.1232 -3.7836  1.4368  0.8784 -1.4526
0.0459
 -0.2871  1.0858  1.7878 -0.7866  0.8229 -1.6191 -1.2861  0.6178 -
0.2012
  0.5008 -0.9033  1.5538 -4.2831  0.8518  0.1498 -1.9953 -1.3713 -
2.9314
 -1.2153 -0.7473 -0.5523 -1.0593  0.4934 -1.7856 -1.9521 -5.6151
1.5928
 -0.6693 -1.8855 -2.4633 -2.6193  0.7348 -0.8865 -1.6833  1.3198 -
5.1156
 -0.3183  0.3731 -1.2543 -0.7993 -2.1903  0.7816  0.6009  0.3002 -
3.6171
 -2.6181 -1.5273 -2.7754  1.5081 -2.1513 -2.2851  1.6708 -0.7863 -
2.3073
 -3.4506 -0.3684  0.7135 -0.6867 -3.0874 -5.7816  0.3838  0.3123
0.1125
 -1.4859 -0.0873 -3.1176  0.4488  0.8993 -3.1264 -3.0094 -2.0733 -
3.9501
 -2.9511 -1.7223 -0.1374 -4.7826 -1.0905 -0.6504  0.0875 -3.8284 -
1.4883
  0.9117  1.5293  0.6211 -2.7846 -0.1334  0.4395 -1.0203  0.9028 -
1.7056
 -3.2434  0.5121  0.8351  0.0762 -1.6313 -2.4243 -1.1373 -3.3994 -
2.6583
 -0.9884  1.3476 -0.4778 -0.0651 -0.832 ]
******************** extraversion ********************
[ 0.5269  1.2396  0.1637 -0.344  -1.0697 -2.2954 -1.0379  0.01   -
0.6048
 -0.9122  0.0914  0.8171 -0.7585 -0.598   0.672   0.7785 -1.0659
1.3933
 -0.2714 -1.3599 -1.9881  0.1357 -0.9245 -1.7954  0.0552 -0.0537
1.0859
  0.3174  2.1129  0.4711  0.6248  0.2366 -0.5349 -0.4511 -0.6343 -
0.7794
  0.3817  1.8331 -2.2308 -0.6582 -0.2974 -2.6028 -2.4491 -1.2196
2.1617
```

```
 -0.1988 -0.4891  0.9322 -1.2148  1.7007  1.1437 -0.1437  1.8543
1.547
  0.8809 -1.6807  0.7083  0.5994  1.1074 -1.5776  0.9623 -1.8344 -
1.5051
 -0.1626  0.926  -3.2176 -1.3733  1.2525 -1.4688 -2.1418  1.688
1.1558
 -1.6502  0.4906  1.5428 -1.527   1.9782 -1.9405  0.2113 -2.3759
0.2075
 -2.0856 -0.3803  0.2729 -4.6009 -1.2511 -0.0319  2.3154 -3.525  -
0.8157
 -0.6355  0.6984  1.3977 -3.0639  2.1234  2.008  -2.3396 -1.9042
1.1804
 -2.775   0.73    0.164  -0.824  -2.1219 -0.7068  1.4702  1.4267 -
1.1422
 -1.6865 -0.2882 -2.6662  1.0348 -1.0334 -2.9565 -4.2935 -2.7565
0.065
 -0.1996 -3.537  -2.0131 -0.4226 -3.8636  1.5791 -0.1408 -1.9408 -
0.7026
 -2.4485 -2.9102 -0.6339  1.6484 -3.3713  0.3034  1.3614  2.5354 -
0.0933
  1.9801  0.3292 -0.8703 -1.0116 -0.4899  0.1138 -1.7086  0.6388 -
3.9861
 -2.521  -2.8113  0.9042  0.2477 -3.8324 -0.3149 -0.2516 -4.4472 -
1.7083
 -1.2056]
********** nueroticism **********
[ 1.3549  -0.1076  -0.8682  -0.4078   0.09163 -0.7415  -2.0092
0.1459
  0.9066   0.1798  -0.995   -0.2902  -0.6147  -1.6289  -0.2344
0.06223
  0.7798  -0.4879   0.5323   1.8249  -0.3612   0.0623  -1.2303  -
1.5021
  1.1601  -1.8824  -0.735    0.2727  -2.1998  -1.2486  -1.1218
0.2973
  1.7074   0.26793 -2.2879   1.0024   0.653   -0.4821  -0.6428  -0.349
  1.5404   0.3995   0.6498   1.794    0.0192   1.0333   0.4148
0.8848
 -0.7603   0.88483 -0.5253   1.2869   1.4724  -0.8778  -0.1727
0.0035
 -1.1128   1.6672   0.64983  1.1199  -2.3895  -1.7556  -2.136
0.53233
 -0.0552  -0.26087  0.17983 -0.6134   0.219    0.5262   0.76733
0.00353
 -1.3753  -0.29027  1.0611  -0.7015   2.6475   0.7673   1.4136
0.29733
  0.0917  -1.4653   0.7967   2.2949  -1.3478  -0.2609   1.5899  -
0.05527
  2.1774   1.4297  -0.3414  -0.78967 -1.1422   2.301   -0.9953
0.9169
```

```
 -0.52527 -0.87777 -2.643   -0.40777 -0.01    -0.4371   1.58993
2.4278
  1.3255   2.9349   0.6204  -2.0529  -1.8179  -1.05407  2.0599
2.0475
 -1.3184   1.85433 -0.5644  -1.11277 -0.365   -1.4066   1.2374  -
1.9033
  1.9424  -2.5163   0.70853 -1.5828   1.67803  1.7662  -1.58287  2.53
 -0.70157  0.44423  0.4442   2.1743   1.9207   0.3561  -1.7004  -
2.5047
  1.1492   0.3756  -1.671    0.8457  -0.17277 -0.99527 -0.76027 -
1.34787
 -1.23027  0.7086   1.5018  -1.3255   0.973   -2.1704   1.06113
2.1187
  0.40413 -0.0846  -1.9354  -0.43717  1.41363  3.3525   0.4041
1.76613
  1.14923 -2.2627  -0.64277 -1.46537  1.32553  0.1477  -0.7897  -
0.61337
  1.11983  3.235    0.2679   1.70743 -0.34897  2.4125   1.35483
0.62043
  2.6814  -1.0541   1.8543  -1.1911   0.41483 -0.9659   3.0617
1.23733
 -0.7496   2.765    1.82493  2.7356  -1.49467  2.4712  -0.5382
2.29493
 -0.5958   2.47123 -1.5899  -0.08457 -1.7591   3.3152   0.35603
2.2068
 -1.3008  -0.96597 -0.8177   0.7493   1.678    1.00233  1.50173
0.97293
  1.0747   1.5578   2.5546  -1.14217  0.6605   2.5593   0.9553
0.2759
  2.0306 ]
******************** openess_to_experience ********************
[-0.4455  0.8637  0.6721 -0.9194 -0.1295 -0.8608 -1.0872  1.247  -
0.2859
  0.0973  0.0284 -1.2354  1.2528  1.4386  0.3444 -1.3539 -2.7769 -
5.0763
 -0.6692 -0.2875  1.1343 -0.0943 -0.7615  0.2889 -0.4776  1.0554
1.8224
  0.6603 -1.4356 -1.359  -3.1602  0.1864  0.5024 -1.244  -0.1543 -
0.6035
 -5.477  -1.8189 -2.3937  0.3049  0.8183  0.4805  1.6302 -2.2021
1.2923
  0.9763 -2.9731 -1.0524 -1.6273  1.6082 -5.2679 -0.169  -1.0774 -
3.9605
 -0.0506 -0.5245 -1.8673 -0.8799 -0.9984  0.5419 -4.5015 -2.1833 -
2.3415
  0.1275 -0.643  -1.3934  1.4502 -6.9925 -1.7093 -0.8782 -3.4471 -
2.0105
 -0.4137  0.4234 -1.5513 -1.1169 -1.425   0.8973  0.0916 -2.5853 -
2.0253
```

```
  0.0679  0.7788 -0.4139 -3.3518 -3.735  -2.3412 -2.9686  1.0031 -
4.3099
 -0.0844  1.0158 -2.7595  1.3976 -0.406   1.4186 -0.4601 -1.0458 -
5.8428
 -2.0648  0.167   0.7941  1.2121 -3.5434 -6.8009  0.1187  0.585  -
3.9266
 -2.6572 -3.6051 -3.763  -5.686  -1.1291 -0.0167 -2.8152 -1.8278 -
5.6512
  0.7631 -5.4595  1.0395 -0.4392 -3.1311 -0.5081 -1.4724 -2.3017 -
1.6662
 -1.9463  0.9404 -2.4202 -0.2511  0.7906 -1.9234 -6.6092 -7.3757 -
0.1521
  1.4003 -0.8045  0.376  -1.8386  0.7657  0.7104 -0.4229]
```

```
import datetime as dt
dataset["DOJ"]=pd.to_datetime(dataset["DOJ"]).dt.date
dataset["DOL"].replace("present",dt.datetime.today(),inplace=True)
dataset['DOL'] = pd.to_datetime(dataset['DOL']).dt.date
## Converting feature from DOJ and DOL as we are only concerned with
how many years the person has worked in the organ
dataset['Period'] = pd.to_datetime(dataset["DOL"]).dt.year -
pd.to_datetime(dataset['DOJ']).dt.year
## Converting DOB column from timestamp to year
dataset['DOB'] = pd.to_datetime(dataset['DOB']).dt.year
dataset.head(5)
```


```
{"type":"dataframe","variable_name":"dataset"}
```


```
#Graduation year contains 0 value,we need to impute it with mode.
dataset['GraduationYear'].replace(0,dataset.GraduationYear.mode()
[0],inplace=True)
dataset['GraduationYear']=pd.to_datetime(dataset['GraduationYear'])
dataset['gyear']=dataset['GraduationYear'].dt.year
#New columns
dataset['12GradAge']=abs(dataset['12graduation']-dataset['DOB'])
dataset['GradAge']=abs(dataset['gyear']-dataset['DOB'])

# no of 0's per column
(dataset==0).astype(int).sum(axis=0)
```

```
ID                          0
Salary                      0
DOJ                         0
DOL                         0
Designation                 0
JobCity                     0
Gender                      0
DOB                         0
10percentage                0
10board                   350
```

```
12graduation                0
12percentage                0
12board                   359
CollegeID                   0
CollegeTier                 0
Degree                      0
Specialization              0
collegeGPA                  0
CollegeCityID               0
CollegeCityTier          2797
CollegeState                0
GraduationYear              0
English                     0
Logical                     0
Quant                       0
Domain                      0
ComputerProgramming         0
ElectronicsAndSemicon       0
ComputerScience             0
MechanicalEngg              0
ElectricalEngg              0
TelecomEngg                 0
CivilEngg                   0
conscientiousness           0
agreeableness               0
extraversion                0
nueroticism                 0
openess_to_experience       0
Period                    319
gyear                       0
12GradAge                   0
GradAge                     0
dtype: int64
```

```
dataset.isin([-1, 'NaN']).sum()
```

```
ID                          0
Salary                      0
DOJ                         0
DOL                         0
Designation                 0
JobCity                   461
Gender                      0
DOB                         0
10percentage                0
10board                     0
12graduation                0
12percentage                0
12board                     0
CollegeID                   0
```

```
CollegeTier                  0
Degree                       0
Specialization               0
collegeGPA                   0
CollegeCityID                0
CollegeCityTier              0
CollegeState                 0
GraduationYear               0
English                      0
Logical                      0
Quant                        0
Domain                     246
ComputerProgramming        868
ElectronicsAndSemicon     2854
ComputerScience           3096
MechanicalEngg            3763
ElectricalEngg            3837
TelecomEngg               3624
CivilEngg                 3956
conscientiousness            0
agreeableness                0
extraversion                 0
nueroticism                  0
openess_to_experience        0
Period                       0
gyear                        0
12GradAge                    0
GradAge                      0
dtype: int64
```

```
dataset[dataset["Designation"]=="get"]
[['Designation','JobCity','Salary','Specialization']]
```


```
{"summary":"{\n  \"name\":
\"dataset[dataset[\\\"Designation\\\"]==\\\"get\\\"]
[['Designation','JobCity','Salary','Specialization']]\",\n  \"rows\":
14,\n  \"fields\": [\n    {\n      \"column\": \"Designation\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 1,\n        \"samples\": [\n          \"get\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"JobCity\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 12,\n        \"samples\": [\n
\"MEERUT\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"Salary\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 287515,\n        \"min\": 110000,\n        \"max\": 1210000,\
n        \"num_unique_values\": 12,\n        \"samples\": [\n
280000\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
```

```
\"Specialization\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 5,\n        \"samples\":
[\n          \"mechanical and automation\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe"}
```

```
# For people with mechanical engineering or mechanical and automation
specialization
mech = dataset[dataset['Specialization'].isin(['mechanical
engineering', 'mechanical and automation'])]['Designation'].mode()[0]

# For people with electronics and electrical engineering
specialization
eee = dataset[dataset['Specialization'] == 'electronics and electrical
engineering']['Designation'].mode()[0]

print(f'Mode for mechanical: {mech}\nMode for EEE: {eee}')
```

```
Mode for mechanical: production engineer
Mode for EEE: system engineer
```

```
# For mechanical domain
dataset.loc[dataset['Specialization'].isin(['mechanical engineering',
'mechanical and automation']), 'Designation'] =
dataset.loc[dataset['Specialization'].isin(['mechanical engineering',
'mechanical and automation']), 'Designation'].replace('old_value',
'new_value')
# For EEE domain
dataset['Designation'].replace('old_value', eee, inplace=True)
```

```
dataset['JobCity'].replace(-1,'unknown',inplace=True)
dataset['JobCity'].apply(lambda x:x.title())
```

```
0               Bangalore
1                  Indore
2                 Chennai
3                 Gurgaon
4                 Manesar
               ...
3993           New Delhi
3994            Hyderabad
3995            Bangalore
3996    Asifabadbanglore
3997              Chennai
Name: JobCity, Length: 3998, dtype: object
```

```
dataset[dataset["JobCity"]=='unknown']
```

```
{"type":"dataframe"}
```

```
dataset[dataset["JobCity"]=="unknown"]
[["Designation","12GradAge","GradAge","JobCity","Gender","10percentage
","10board"]]
```


```
{"summary":"{\n  \"name\":
\"dataset[dataset[\\\"JobCity\\\"]==\\\"unknown\\\"]
[[\\\"Designation\\\",\\\"12GradAge\\\",\\\"GradAge\\\",\\\"JobCity\\\
",\\\"Gender\\\",\\\"10percentage\\\",\\\"10board\\\"]]\",\n
\"rows\": 461,\n  \"fields\": [\n    {\n      \"column\":
\"Designation\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 155,\n
\"samples\": [\n          \"senior systems engineer\",\n
\"executive assistant\",\n          \"get\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"12GradAge\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
0,\n        \"min\": 15,\n        \"max\": 21,\n
\"num_unique_values\": 7,\n        \"samples\": [\n          18,\n
16,\n          20\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"GradAge\",\n      \"properties\": {\n        \"dtype\": \"number\",\
n        \"std\": 1,\n        \"min\": 14,\n        \"max\": 24,\n
\"num_unique_values\": 11,\n        \"samples\": [\n          19,\n
20,\n          15\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"JobCity\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 1,\n        \"samples\":
[\n          \"unknown\"\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    },\n    {\n
\"column\": \"Gender\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          \"m\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"10percentage\",\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": 9.660928712872739,\n        \"min\":
45.33,\n        \"max\": 95.52,\n        \"num_unique_values\": 250,\n
\"samples\": [\n          92.12\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"10board\",\n      \"properties\":
{\n        \"dtype\": \"category\",\n        \"num_unique_values\":
54,\n        \"samples\": [\n          \"board of secondary
education\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    }\n  ]\n}","type":"dataframe"}
```


```
### cleaning the column which have similar meaning
dataset["JobCity"].replace("Bangalore","Bengaluru",inplace=True)
dataset["JobCity"].replace("Banaglore","Bengaluru",inplace=True)
dataset["JobCity"].replace("Chennai,
Bangalore","Bengaluru",inplace=True)
dataset["JobCity"].replace(" Bangalore","Bengaluru",inplace=True)
```

```
dataset["JobCity"].replace("Bangalore ","Bengaluru",inplace=True)
dataset["JobCity"].replace("Banglore","Bengaluru",inplace=True)
dataset["JobCity"].replace("Jaipur ","Jaipur",inplace=True)
dataset["JobCity"].replace("Gandhinagar","Gandhi Nagar",inplace=True)
dataset["JobCity"].replace("Bangalore ","Bengaluru",inplace=True)
dataset["JobCity"].replace("Jaipur ","Jaipur",inplace=True)
dataset["JobCity"].replace("Gandhinagar","Gandhi Nagar",inplace=True)
dataset["JobCity"].replace("Hyderabad ","Hyderabad",inplace=True)
dataset["JobCity"].replace("Hyderabad(Bhadurpally)","Hyderabad",inplac
e=True)
dataset["JobCity"].replace("Bhubaneswar ","Bhubaneswar",inplace=True)
dataset["JobCity"].replace("Delhi/Ncr","Delhi",inplace=True)
dataset["JobCity"].replace("Nagpur ","Nagpur",inplace=True)
dataset["JobCity"].replace("Pune ","Pune",inplace=True)
dataset["JobCity"].replace("Trivandrum ","Trivandrum",inplace=True)
dataset["JobCity"].replace("Thiruvananthapuram","Trivandrum",inplace=T
rue)

best_mode = []
best_mode.append(dataset[dataset["Designation"]=="software engineer"]
["JobCity"].mode().to_list()[0])
best_mode.append(dataset[dataset["Gender"]=="m"]
["JobCity"].mode().to_list()[0])
best_mode.append(dataset[dataset["10percentage"]==76]
["JobCity"].mode().to_list()[0])
best_mode.append(dataset[dataset["10board"]=="cbse"]
["JobCity"].mode().to_list()[0])
best_mode.append(dataset[dataset["12percentage"]==64]
["JobCity"].mode().to_list()[0])
best_mode.append(dataset[dataset["12board"]=="cbse"]
["JobCity"].mode().to_list()[0])
best_mode.append(dataset[dataset["collegeGPA"]==70]
["JobCity"].mode().to_list()[0])
best_mode.append(dataset[dataset["Salary"]==200000]
["JobCity"].mode().to_list()[0])
best_mode.append(dataset[dataset["Degree"].str.startswith("B.Tech/")]
["JobCity"].mode().to_list()[0])
best_mode.append(dataset[dataset["Specialization"].str.startswith("ele
ctronics and communication eng")]["JobCity"].mode().to_list()[0])
best_mode.append(dataset[dataset["CollegeState"].str.startswith("Uttar
Pradesh")]["JobCity"].mode().to_list()[0])
best_mode
```

```
['Bengaluru',
 'Bengaluru',
 'unknown',
 'Noida',
 'unknown',
 'Noida',
 'Bengaluru',
```

```
 'unknown',
 'Bengaluru',
 'Bengaluru',
 'Noida']
```

```
### We can see mode from the best_mode list is 'Bangalore'
dataset["JobCity"].replace("unknown",'Bengaluru',inplace=True)
```

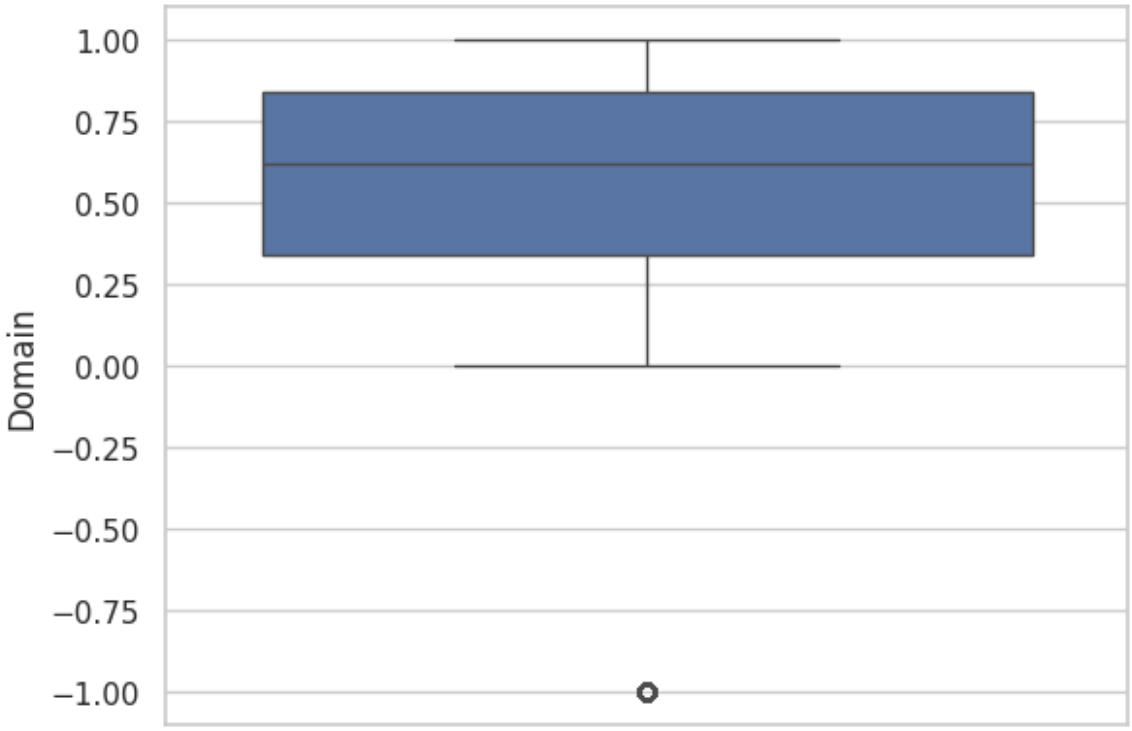
```
dataset[dataset["10board"] == 0][["Designation", "12GradAge",
"GradAge", "JobCity", "Gender", "10percentage", "10board",
"12percentage","12board","collegeGPA","Salary","Degree","Specializatio
n","CollegeState"]]
```

```
{"repr_error":"'str' object has no attribute
'empty'","type":"dataframe"}
```

```
best_value2=[]
best_value2.append(dataset[dataset["Designation"]=="software
engineer"]["10board"].mode().to_list()[0])
best_value2.append(dataset[dataset["Gender"]=="m"]
["10board"].mode().to_list()[0])
best_value2.append(dataset[dataset["10percentage"]==75]
["10board"].mode().to_list()[0])
best_value2.append(dataset[dataset["JobCity"]=="Bengaluru"]
["10board"].mode().to_list()[0])
best_value2.append(dataset[dataset["12percentage"]==65]
["10board"].mode().to_list()[0])
best_value2.append(dataset[dataset["collegeGPA"]==65]
["10board"].mode().to_list()[0])
best_value2.append(dataset[dataset["Salary"]==400000]
["10board"].mode().to_list()[0])
best_value2.append(dataset[dataset["Degree"].str.startswith("B.Tech/")
]["10board"].mode().to_list()[0])
best_value2.append(dataset[dataset["Specialization"].str.startswith("c
omputer eng")]["10board"].mode().to_list()[0])
best_value2.append(dataset[dataset["CollegeState"].str.startswith("Tam
il Nadu")]["10board"].mode().to_list()[0])
best_value2
```

```
['cbse',
 'cbse',
 'cbse',
 'cbse',
 'cbse',
 'cbse',
 'cbse',
 'cbse',
 'cbse',
 'state board']
```

```
### Replacing with the mode of the best_value list
dataset['10board'].replace(0,'cbse',inplace=True)

dataset['12board'].replace(0,'cbse',inplace=True)

sns.boxplot(dataset['Domain'])
plt.show()
```




```
## As we can see outlier,it is better to use median to replace the
missing values.
dataset['Domain'].replace(-1,dataset['Domain'].median(),inplace=True)
dataset.head()

{"type":"dataframe","variable_name":"dataset"}

replace_list_state=['board of intermediate education,ap', 'state
board',
 'mp board', 'karnataka pre university board', 'up',
 'p u board, karnataka', 'dept of pre-university education', 'bie',
 'kerala state hse board', 'up board', 'bseb', 'chse', 'puc',
 ' upboard',
 'state board of intermediate education, andhra pradesh',
 'karnataka state board',
 'west bengal state council of technical education', 'wbchse',
 'maharashtra state board', 'ssc',
 'sda matric higher secondary school', 'uttar pradesh board', 'ibe',
 'chsc', 'board of intermediate', 'upboard', 'sbtet',
```

```
'hisher seconadry examination(state board)', 'pre university',
'borad of intermediate', 'j & k board',
'intermediate board of andhra pardesh', 'rbse',
'central board of secondary education', 'jkbose', 'hbse',
'board of intermediate education', 'state', 'ms board', 'pue',
'intermediate state board', 'stateboard', 'hsc',
'electonincs and communication(dote)', 'karnataka pu board',
'government polytechnic mumbai , mumbai board', 'pu board',
'baord of intermediate education', 'apbie', 'andhra board',
'tamilnadu stateboard',
'west bengal council of higher secondary education',
'cbse,new delhi', 'u p board', 'intermediate', 'biec,patna',
'diploma in engg (e &tc) tilak maharashtra vidayapeeth',
'hsc pune', 'pu board karnataka', 'kerala', 'gsheb',
'up(allahabad)', 'nagpur', 'st joseph hr sec school',
'pre university board', 'ipe', 'maharashtra', 'kea', 'apsb',
'himachal pradesh board of school education', 'staae board',
'international baccalaureate (ib) diploma', 'nios',
'karnataka board of university',
'board of secondary education rajasthan', 'uttarakhand board',
'ua', 'scte vt orissa', 'matriculation',
'department of pre-university education', 'wbscte',
'preuniversity board(karnataka)', 'jharkhand accademic council',
'bieap', 'msbte (diploma in computer technology)',
'jharkhand acamedic council (ranchi)',
'department of pre-university eduction', 'biec',
'sjrcw', ' board of intermediate', 'msbte',
'sri sankara vidyalaya', 'chse, odisha', 'bihar board',
'maharashtra state(latur board)', 'rajasthan board', 'mpboard',
'state board of technical eduction panchkula', 'upbhsie', 'apbsc',
'state board of technical education and training',
'secondary board of rajasthan',
'tamilnadu higher secondary education board',
'jharkhand academic council',
'board of intermediate education,hyderabad', 'up baord', 'pu',
'dte', 'board of secondary education', 'pre-university',
'board of intermediate education,andhra pradesh',
'up board , allahabad', 'srv girls higher sec school,rasipuram',
'intermediate board of education,andhra pradesh',
'intermediate board examination',
'department of pre-university education, bangalore',
'stmiras college for girls', 'mbose',
'department of pre-university education(government of karnataka)',
'dpue', 'msbte pune', 'board of school education harayana',
'sbte, jharkhand', 'bihar intermediate education council, patna',
'higher secondary', 's j polytechnic', 'latur',
'board of secondary education, rajasthan', 'jyoti nivas', 'pseb',
'biec-patna', 'board of intermediate education,andra pradesh',
'chse,orissa', 'pre-university board', 'mp', 'intermediate board',
```

```
'govt of karnataka department of pre-university education',
'karnataka education board',
'board of secondary school of education', 'pu board ,karnataka',
'karnataka secondary education board', 'karnataka sslc',
'board of intermediate ap', 'u p', 'state board of karnataka',
'directorate of technical education,banglore', 'matric board',
'andhpradesh board of intermediate education',
'stjoseph of cluny matrhrsecschool,neyveli,cuddalore district',
'bte up', 'scte and vt ,orissa', 'hbsc',
'jawahar higher secondary school', 'nagpur board', 'bsemp',
'board of intermediate education, andhra pradesh',
'board of higher secondary orissa',
'board of secondary education,rajasthan(rbse)',
'board of intermediate education:ap,hyderabad', 'science college',
'karnatak pu board', 'aissce', 'pre university board of karnataka',
'bihar', 'kerala state board', 'uo board',
'karnataka board', 'tn state board',
'kolhapur divisional board, maharashtra',
'jaycee matriculation school',
'board of higher secondary examination, kerala',
'uttaranchal state board', 'intermidiate', 'bciec,patna', 'bice',
'karnataka state', 'state broad', 'wbbhse', 'gseb',
'uttar pradesh', 'ghseb', 'board of school education uttarakhand',
'gseb/technical education board', 'msbshse,pune',
'tamilnadu state board', 'board of technical education',
'kerala university', 'uttaranchal shiksha avam pariksha parishad',
'chse(concil of higher secondary education)',
'bright way college, (up board)', 'board of intermidiate',
'higher secondary state certificate', 'karanataka secondary board',
'maharashtra board', 'cgbse', 'diploma in computers', 'bte,delhi',
'rajasthan board ajmer', 'mpbse', 'pune board',
'state board of technical education', 'gshseb',
'amravati divisional board', 'dote (diploma - computer engg)',
'karnataka pre-university board', 'jharkhand board',
'punjab state board of technical education & industrial training',
'department of technical education',
'sri chaitanya junior kalasala', 'state board (jac, ranchi)',
'aligarh muslim university', 'tamil nadu state board', 'hse',
'karnataka secondary education', 'state board ',
'karnataka pre unversity board',
'ks rangasamy institute of technology',
'karnataka board secondary education', 'narayana junior college',
'bteup', 'board of intermediate(bie)', 'hsc maharashtra board',
'tamil nadu state', 'uttrakhand board', 'psbte',
'stateboard/tamil nadu', 'intermediate council patna',
'technical board, punchkula', 'board of intermidiate examination',
'sri kannika parameswari highier secondary school, udumalpet',
'ap board', 'nashik board', 'himachal pradesh board',
'maharashtra satate board',
```

```
 'andhra pradesh board of secondary education',
 'tamil nadu polytechnic',
 'maharashtra state board mumbai divisional board',
 'department of pre university education',
 'dav public school,hehal', 'board of intermediate education, ap',
 'rajasthan board of secondary education',
 'department of technical education, bangalore', 'chse,odisha',
 'maharashtra nasik board',
 'west bengal council of higher secondary examination (wbchse)',
 'holy cross matriculation hr sec school', 'cbsc',
 'pu board karnataka', 'biec patna', 'kolhapur', 'bseb, patna',
 'up board allahabad', 'nagpur board,nagpur', 'diploma(msbte)',
 'dav public school', 'pre university board, karnataka',
 'ssm srsecschool', 'state bord', 'jstb,jharkhand',
 'intermediate board of education', 'mp board bhopal', 'pub',
 'madhya pradesh board', 'bihar intermediate education council',
 'west bengal council of higher secondary eucation',
 'mpc',
 'certificate for higher secondary education (chse)orissa',
 'maharashtra state board for hsc',
 'board of intermeadiate education', 'latur board',
 'andhra pradesh', 'karnataka pre-university',
 'lucknow public college', 'nagpur divisional board',
 'ap intermediate board', 'cgbse raipur', 'uttranchal board',
 'jiec',
 'bihar school examination board patna',
 'state board of technical education harayana', 'mp-bse',
 'up bourd', 'dav public school sec 14',
 'haryana state board of technical education chandigarh',
 'council for indian school certificate examination',
 'jaswant modern school', 'madhya pradesh open school',
 'aurangabad board', 'j&k state board of school education',
 'diploma ( maharashtra state board of technical education)',
 'board of technicaleducation ,delhi',
 'maharashtra state boar of secondary and higher secondary education',
 'hslc (tamil nadu state board)',
 'karnataka state examination board', 'puboard', 'nasik',
 'west bengal board of higher secondary education',
 'up board,allahabad', 'board of intrmediate education,ap',
 'karnataka state pre- university board',
 'state board - west bengal council of higher secondary education :
wbchse',
 'maharashtra state board of secondary & higher secondary education',
 'biec, patna', 'state syllabus', 'cbse board', 'scte&vt',
 'board of intermediate,ap',
 'secnior secondary education board of rajasthan',
 'maharashtra board, pune', 'rbse (state board)',
 'board of intermidiate education,ap',
 'board of high school and intermediate education uttarpradesh',
```

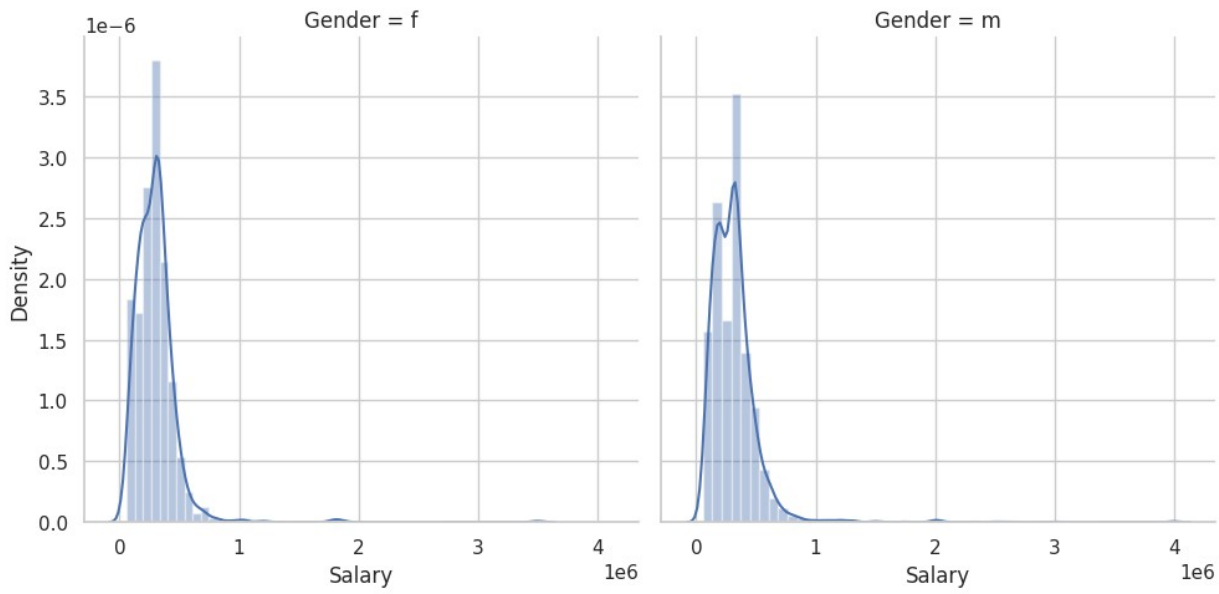
```
 'higher secondary education',
 'board fo intermediate education, ap', 'intermedite',
 'ap board for intermediate education', 'ahsec',
 'punjab state board of technical education & industrial training,
chandigarh',
 'state board - tamilnadu', 'jharkhand acedemic council',
 'scte & vt (diploma)', 'karnataka pu',
 'board of intmediate education ap', 'up-board',
 'boardofintermediate','intermideate','up bord','andhra pradesh state
board','gujarat board']

#replacing the redundant values of the 12board column with
'state','cbse','icse'
for i in replace_list_state:
 dataset['12board'].replace(i,'state',inplace=True)
replace_list_cbse=['cbse',
 'all india board',
 'central board of secondary education, new delhi', 'cbese']
for i in replace_list_cbse:
 dataset['12board'].replace(i,'cbse',inplace=True)
replace_list_icse=[ 'isc', 'icse', 'isc board', 'isce', 'cicse',
 'isc board , new delhi']
for i in replace_list_icse:
 dataset['12board'].replace(i,'icse',inplace=True)
dataset['12board'].unique()

array(['state', 'cbse', 'icse',
       'state  board of intermediate education, andhra pradesh',
       'pu  board karnataka'], dtype=object)

specialization_map= \
{'electronics and communication engineering' : 'EC',
'computer science & engineering' : 'CS',
'information technology' : 'CS' ,
'computer engineering' : 'CS',
'computer application' : 'CS',
'mechanical engineering' : 'ME',
'electronics and electrical engineering' : 'EC',
'electronics & telecommunications' : 'EC',
'electrical engineering' : 'EL',
'electronics & instrumentation eng' : 'EC',
'civil engineering' : 'CE',
'electronics and instrumentation engineering' : 'EC',
'information science engineering' : 'CS',
'instrumentation and control engineering' : 'EC',
'electronics engineering' : 'EC',
'biotechnology' : 'other',
'other' : 'other',
'industrial & production engineering' : 'other',
'chemical engineering' : 'other',
```

```
'applied electronics and instrumentation' : 'EC',
'computer science and technology' : 'CS',
'telecommunication engineering' : 'EC',
'mechanical and automation' : 'ME',
'automobile/automotive engineering' : 'ME',
'instrumentation engineering' : 'EC',
'mechatronics' : 'ME',
'electronics and computer engineering' : 'CS',
'aeronautical engineering' : 'ME',
'computer science' : 'CS',
'metallurgical engineering' : 'other',
'biomedical engineering' : 'other',
'industrial engineering' : 'other',
'information & communication technology' : 'EC',
'electrical and power engineering' : 'EL',
'industrial & management engineering' : 'other',
'computer networking' : 'CS',
'embedded systems technology' : 'EC',
'power systems and automation' : 'EL',
'computer and communication engineering' : 'CS',
'information science' : 'CS',
'internal combustion engine' : 'ME',
'ceramic engineering' : 'other',
'mechanical & production engineering' : 'ME',
'control and instrumentation engineering' : 'EC',
'polymer technology' : 'other',
'electronics' : 'EC'}

dataset['Specialization'] =
dataset['Specialization'].map(specialization_map)
dataset['Specialization'].unique()

array(['CS', 'EC', 'ME', 'EL', 'CE', 'other'], dtype=object)

dataset.drop(columns=['CollegeID','CollegeCityID','CollegeCityTier'],a
xis=1,inplace=True)

dataset.columns

Index(['ID', 'Salary', 'DOJ', 'DOL', 'Designation', 'JobCity',
'Gender', 'DOB',
       '10percentage', '10board', '12graduation', '12percentage',
'12board',
       'CollegeTier', 'Degree', 'Specialization', 'collegeGPA',
'CollegeState',
       'GraduationYear', 'English', 'Logical', 'Quant', 'Domain',
       'ComputerProgramming', 'ElectronicsAndSemicon',
'ComputerScience',
       'MechanicalEngg', 'ElectricalEngg', 'TelecomEngg', 'CivilEngg',
       'conscientiousness', 'agreeableness', 'extraversion',
```

```
'nueroticism',
       'openess_to_experience', 'Period', 'gyear', '12GradAge',
'GradAge'],
      dtype='object')
```

```
# Adjusting salary for individuals with reported salary less than or
equal to 50000
dataset.loc[dataset['Salary'] <= 50000, 'Salary'] *= 12

# List of engineering fields to handle missing or invalid values
fields_to_clean = ['ComputerProgramming', 'ElectronicsAndSemicon',
'ComputerScience', 'MechanicalEngg', 'ElectricalEngg', 'TelecomEngg']

# Replace -1 with 0 in the specified engineering fields
for field in fields_to_clean:
    dataset[field].replace(-1, 0, inplace=True)

plt.figure(figsize=(15,5))
colors = sns.color_palette('bright', n_colors=2)

# Creating a FacetGrid to plot distribution of salaries based on
gender
sns.FacetGrid(dataset, col="Gender", height=5, palette=colors) \
    .map(sns.distplot, "Salary", bins=50) \
    .add_legend()

plt.show()
```

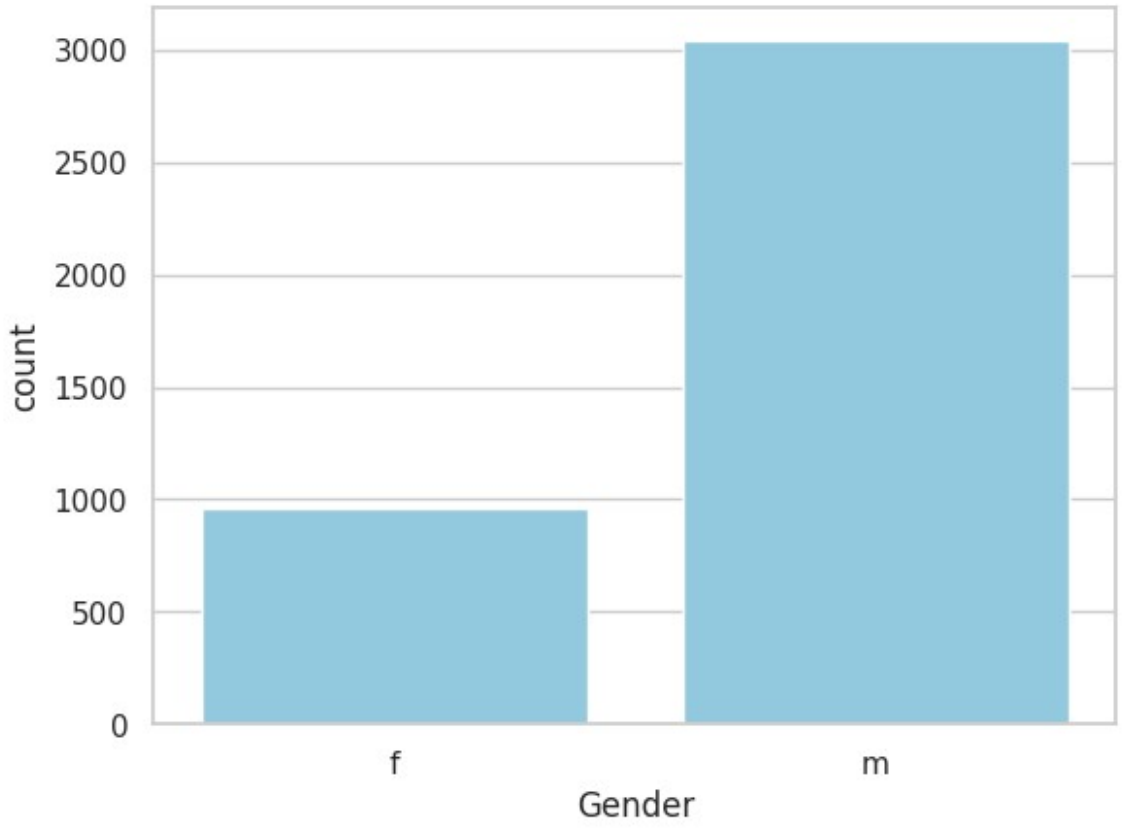
```
<Figure size 1500x500 with 0 Axes>
```



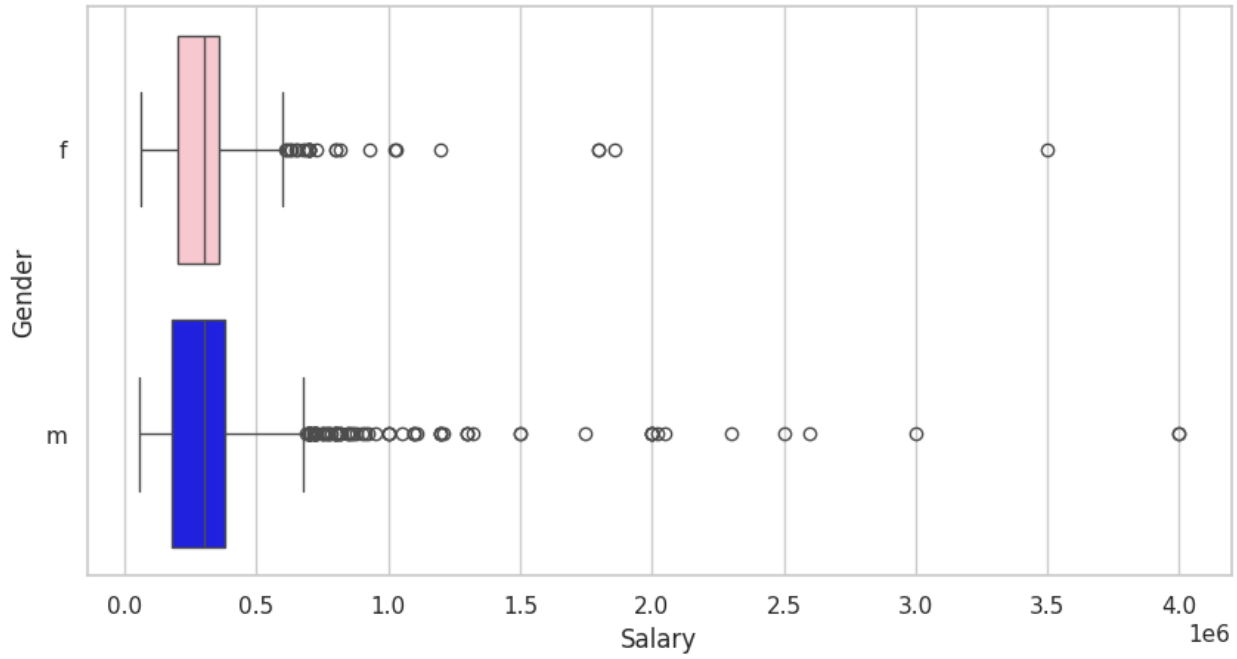

```
bar_color = 'skyblue'
```

```
# Creating a vertical count plot of the 'Gender' column with specified color
sns.countplot(x='Gender', data=dataset, color=bar_color)

# Printing the count of each gender category
print(dataset['Gender'].value_counts())
```

```
m    3041
f     957
Name: Gender, dtype: int64
```




```
plt.figure(figsize=(10,5))

# Specifying colors for the box plot
colors = {'m': 'blue', 'f': 'pink'}

# Creating a horizontal box plot of salary distribution across genders
sns.boxplot(y='Gender', x='Salary', data=dataset, palette=colors)

plt.show()
```




```
plt.figure(figsize=(15,5))

# Define a color palette with different colors for each specialization
palette = sns.color_palette("husl",
n_colors=len(dataset['Specialization'].unique()))

# Creating a box plot of salary distribution across specializations
with specified palette
sns.boxplot(x='Salary', y='Specialization', data=dataset,
palette=palette)

# Adding a title
plt.suptitle('Salary Levels by Specialization')

plt.show()
```


```
### Designation
popular_Designation = dataset['Designation'].value_counts()
[:20].index.tolist()
print(popular_Designation)
```

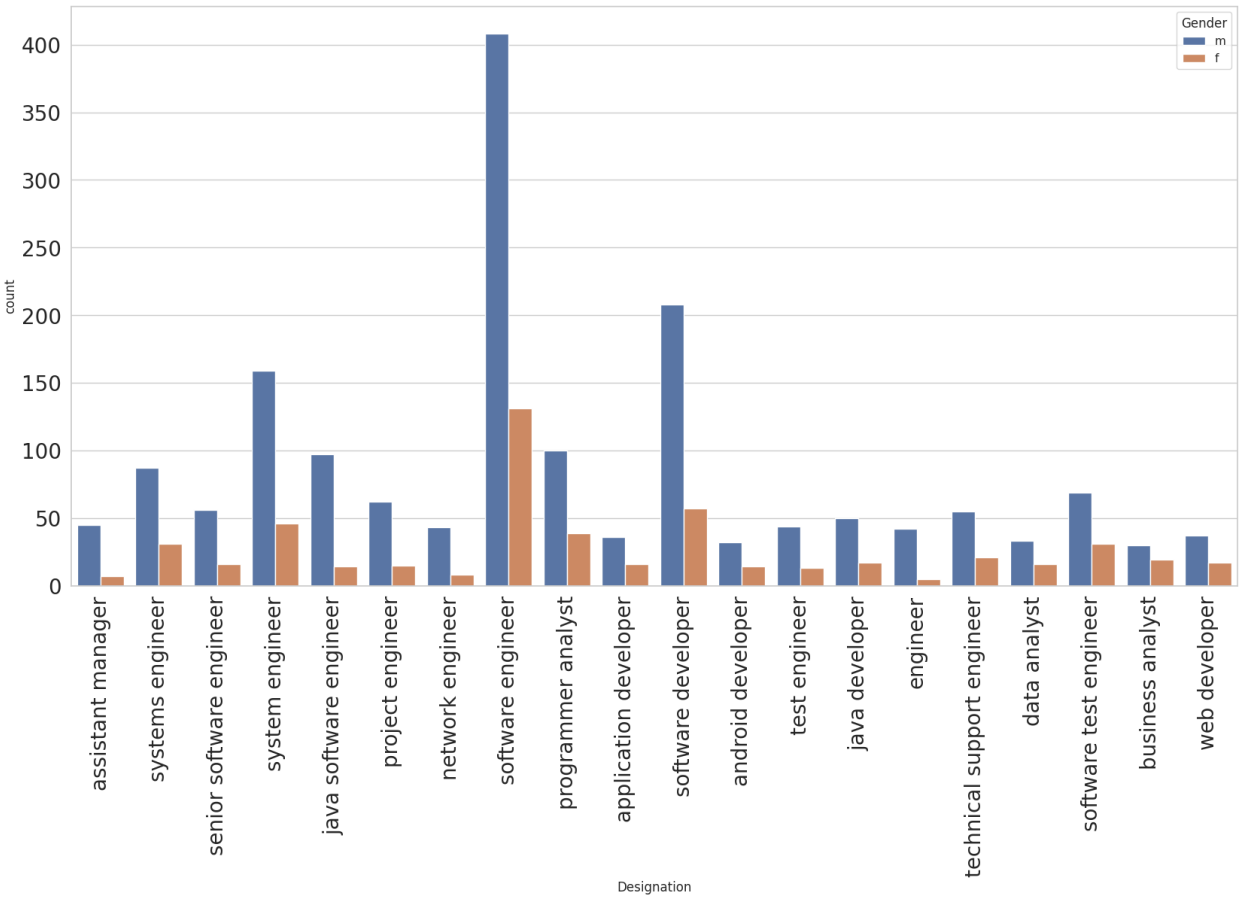
```
['software engineer', 'software developer', 'system engineer',
'programmer analyst', 'systems engineer', 'java software engineer',
'software test engineer', 'project engineer', 'technical support
engineer', 'senior software engineer', 'java developer', 'test
engineer', 'web developer', 'application developer', 'assistant
manager', 'network engineer', 'data analyst', 'business analyst',
'engineer', 'android developer']
```

```
### Unique professions
top_Designations =
dataset[dataset['Designation'].isin(popular_Designation)]
print(f"Unique professions : {len(dataset['Designation'].unique())}")
top_Designations.head()
```

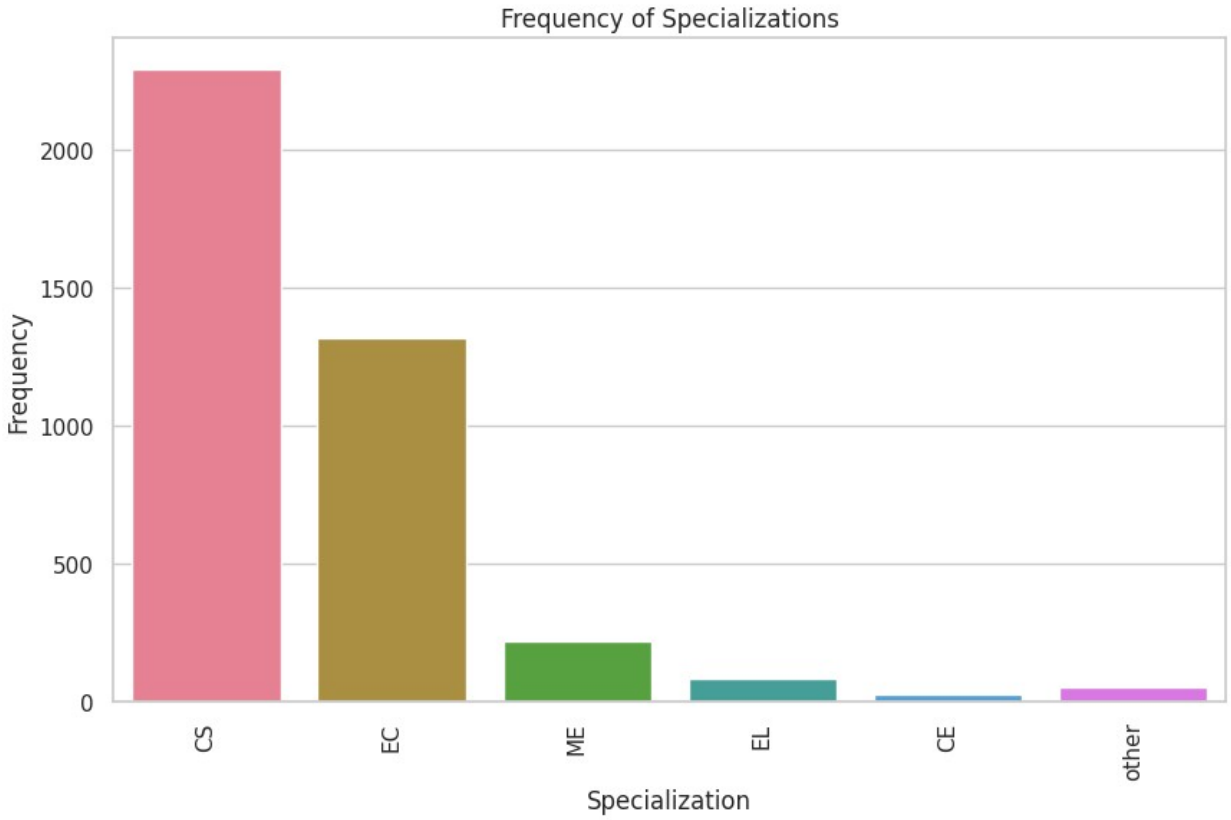
```
Unique professions : 419
```

```
{"type":"dataframe","variable_name":"top_Designations"}
```

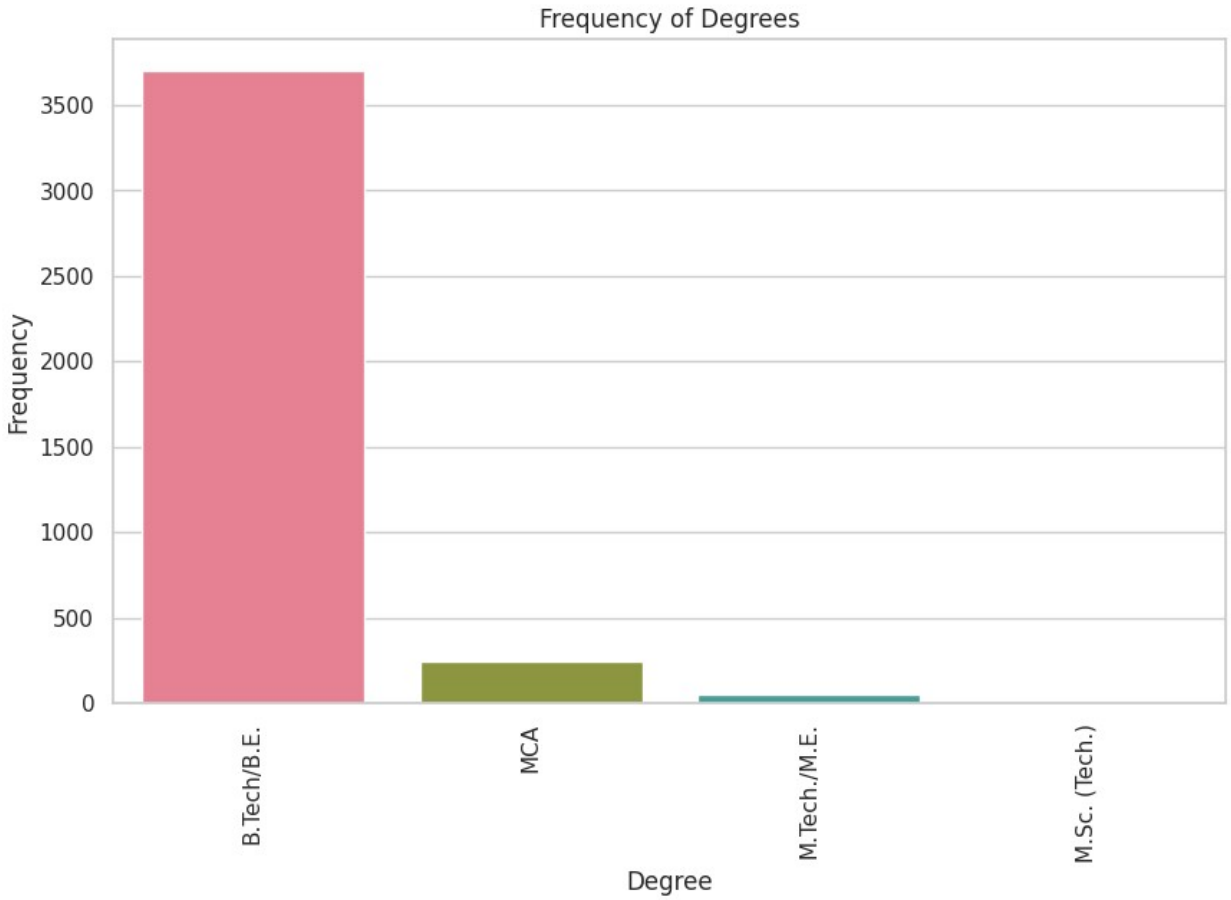
```
plt.figure(figsize=(20,10))
sns.countplot(x='Designation',hue='Gender',data=top_Designations)
plt.xticks(fontsize=20,rotation=90)
plt.yticks(fontsize=20)
plt.show()
```



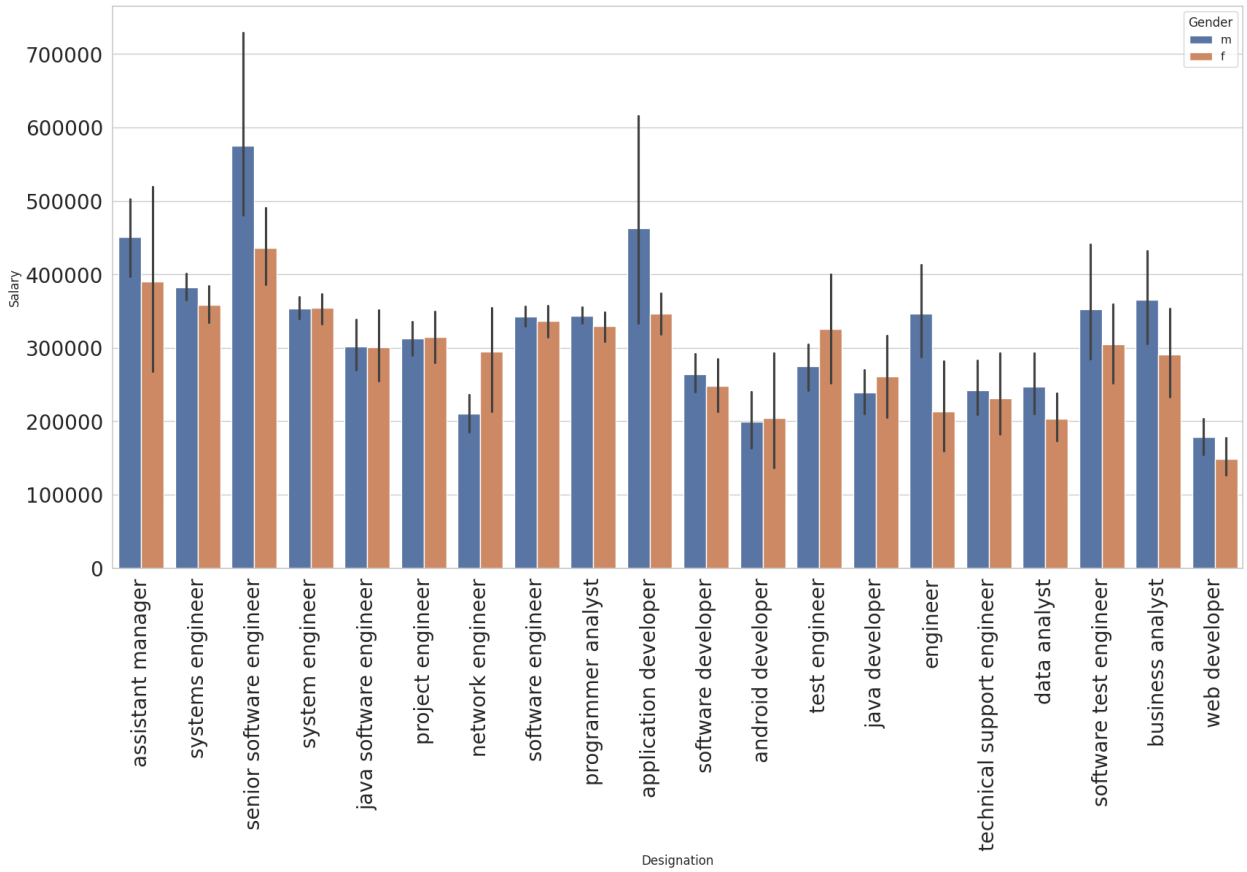

```
plt.figure(figsize=(10, 6))

# Define a color palette with different colors for each specialization
palette = sns.color_palette("husl",
n_colors=len(dataset['Specialization'].unique()))

# Creating a count plot of specialization with specified palette
sns.countplot(x='Specialization', data=dataset, palette=palette)

# Adding titles and labels
plt.title('Frequency of Specializations')
plt.xlabel('Specialization')
plt.ylabel('Frequency')

plt.xticks(rotation=90)  # Rotate x-axis labels for better readability

plt.show()
```



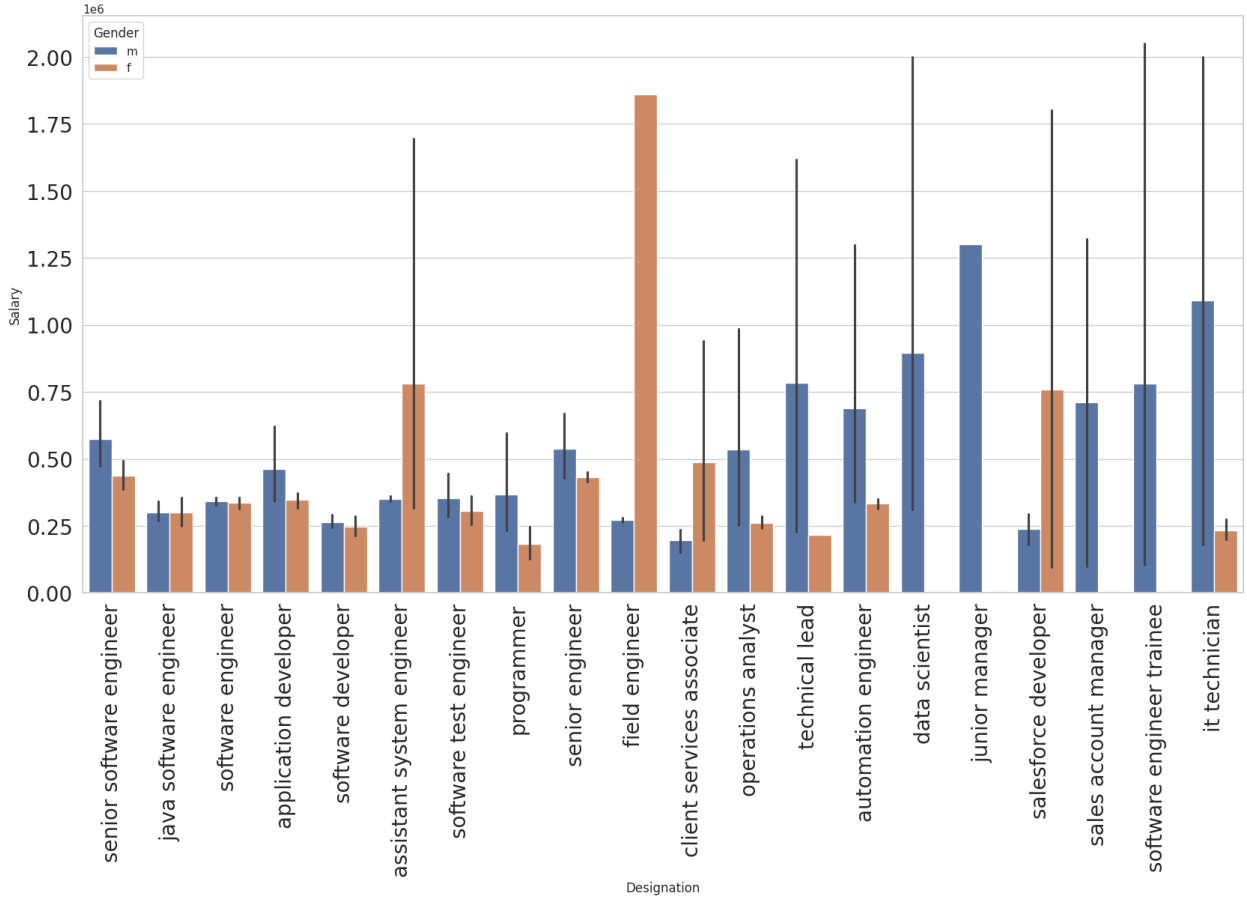

```
plt.figure(figsize=(10, 6))
# Define a color palette with different colors for each degree
palette = sns.color_palette("husl",
n_colors=len(dataset['Degree'].unique()))
# Creating a count plot of degrees with specified palette
sns.countplot(x='Degree', data=dataset, palette=palette)
# Adding titles and labels
plt.title('Frequency of Degrees')
plt.xlabel('Degree')
plt.ylabel('Frequency')
# Rotating x-axis labels for better readability
plt.xticks(rotation=90)
plt.show()
```



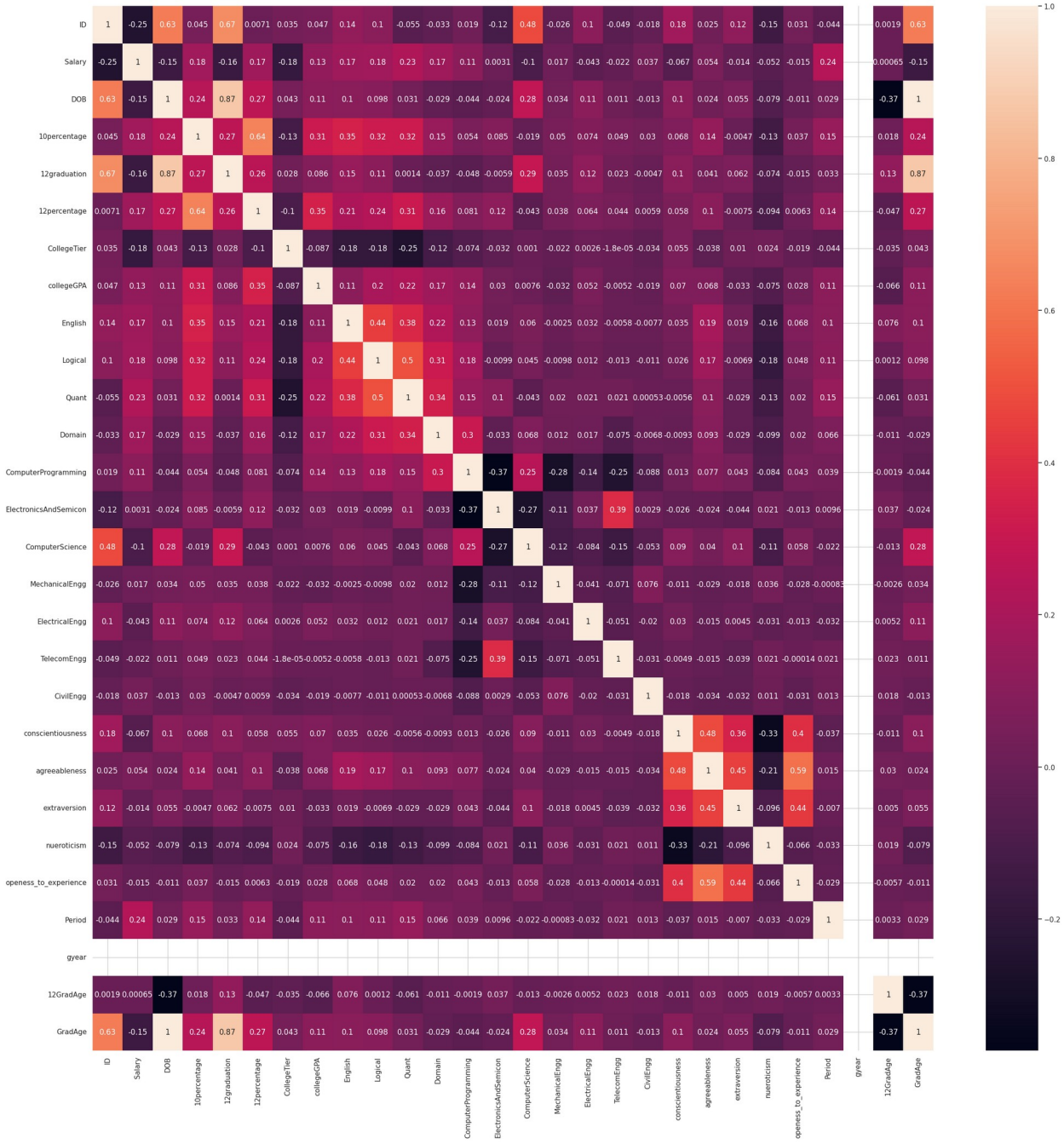

```
plt.figure(figsize=(20,10))
sns.barplot(x='Designation',y='Salary',hue='Gender',data=top_Designati
ons)
plt.xticks(fontsize=20,rotation=90)
plt.yticks(fontsize=20)
plt.show()
```




```
### High paying designations and their relation with respect to gender
high = list(dataset.sort_values("Salary",ascending=False)
["Designation"].unique())[:20]
high_pay = dataset[dataset['Designation'].isin(high)]
high_pay.head()

{"type":"dataframe","variable_name":"high_pay"}

plt.figure(figsize=(20,10))
sns.barplot(x='Designation',y='Salary',hue='Gender',data=high_pay)
plt.xticks(fontsize=20,rotation=90)
plt.yticks(fontsize=20)
plt.show()
```



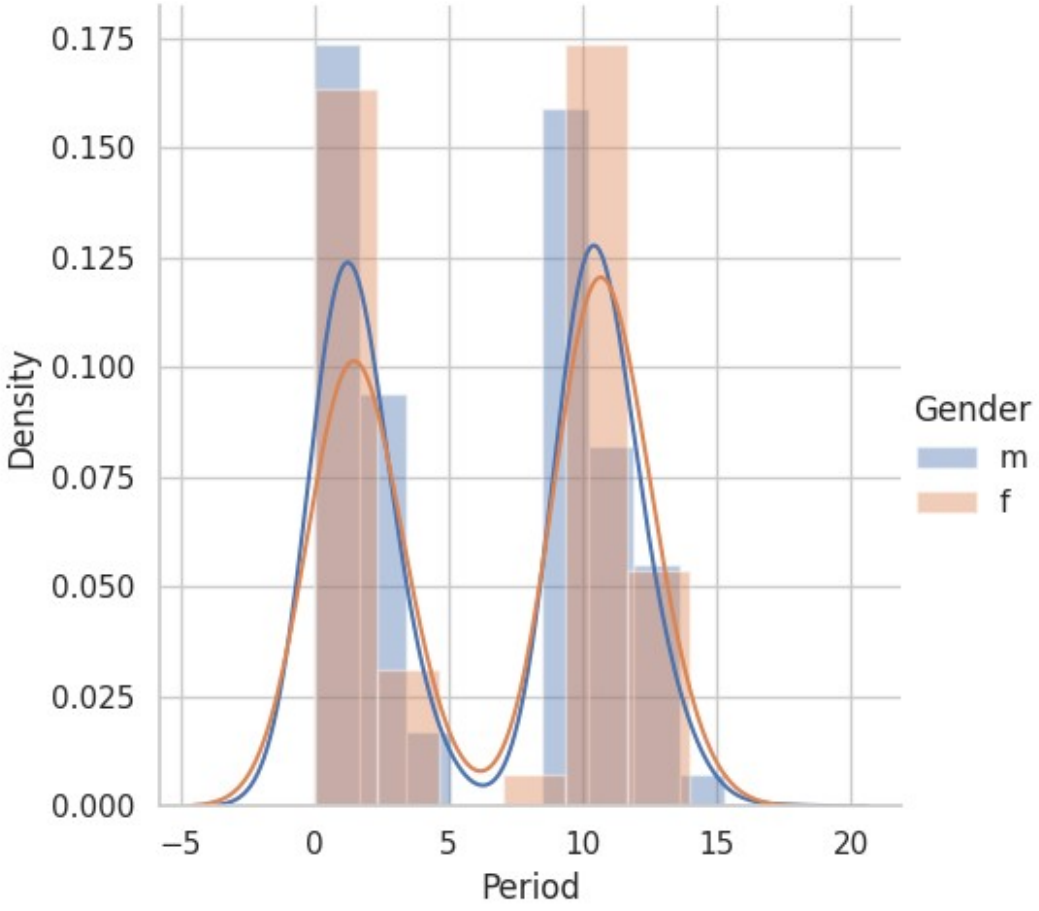

```
plt.figure(figsize=(30,30))

# Creating a heatmap of the correlation matrix with annotations
sns.heatmap(dataset.corr(), annot=True)

plt.show()
```




```
plt.figure(figsize=(20, 5))

# Creating a FacetGrid to visualize the distribution of experience
periods for both genders
sns.FacetGrid(high_pay, hue="Gender", height=5) \
    .map(sns.distplot, "Period") \
    .add_legend()

# Printing mean experience period for the whole dataset
print('For Whole dataset')
```

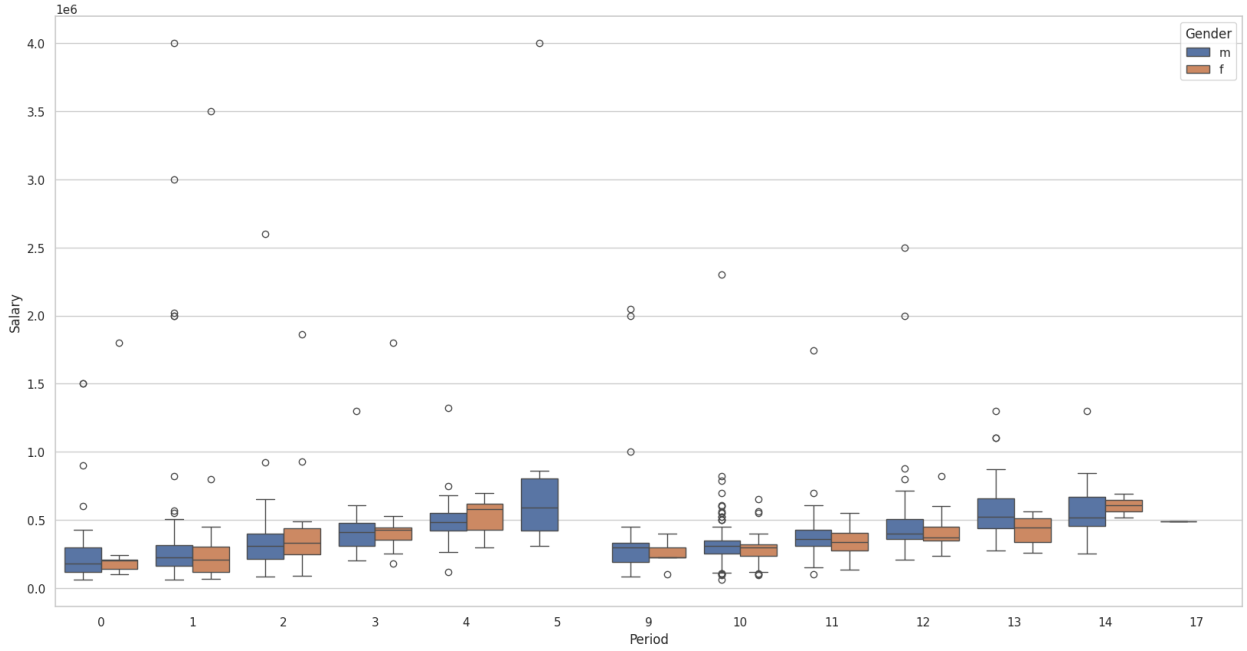
```
print(high_pay.groupby('Gender').Period.mean())
print('*'*20)

# Printing mean experience period for high-paying jobs
print('For High_paying jobs')
print(dataset.groupby('Gender').Period.mean())

plt.show()
```

```
For Whole dataset
Gender
f    6.631579
m    6.239308
Name: Period, dtype: float64
********************
For High_paying jobs
Gender
f    5.809822
m    5.737915
Name: Period, dtype: float64

<Figure size 2000x500 with 0 Axes>
```

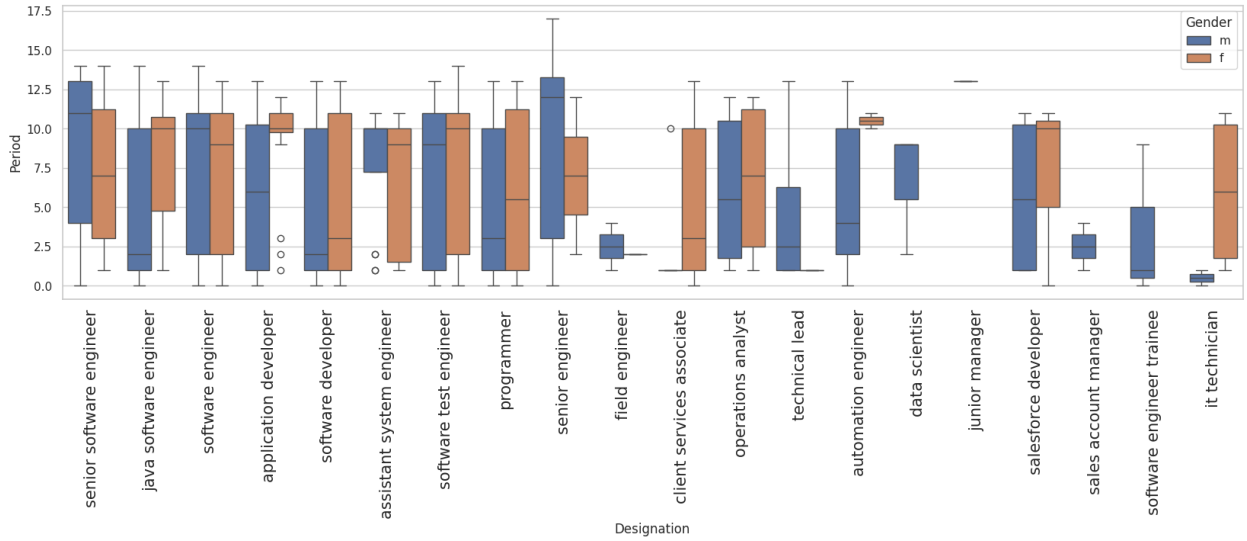


```
plt.figure(figsize=(20,10))
sns.boxplot(data=high_pay,x='Period',y='Salary',hue='Gender')
```

```
<Axes: xlabel='Period', ylabel='Salary'>
```



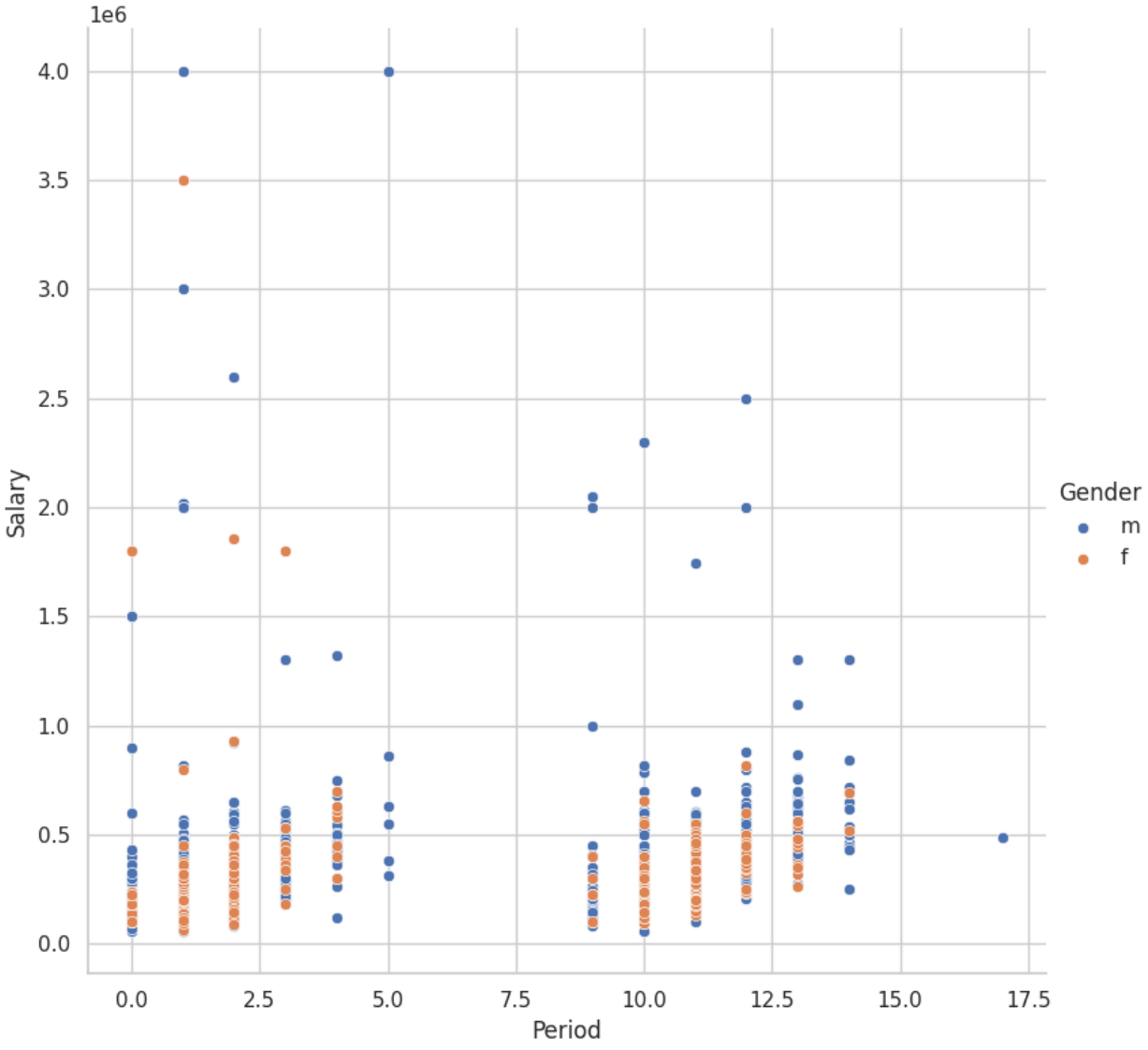

```
plt.figure(figsize=(20,5))
sns.boxplot(data=high_pay,x='Designation',y='Period',hue='Gender')
plt.xticks(fontsize=15,rotation=90)
plt.show()
```




```
plt.figure(figsize=(10, 8))  # Adjust the figure size as needed
```

```
# Creating a FacetGrid to visualize the scatter plot of experience
periods vs. salary for both genders
sns.FacetGrid(data=high_pay, hue='Gender', height=8) \
    .map(sns.scatterplot, 'Period', 'Salary') \
    .add_legend()

plt.show()
```

```
<Figure size 1000x800 with 0 Axes>
```

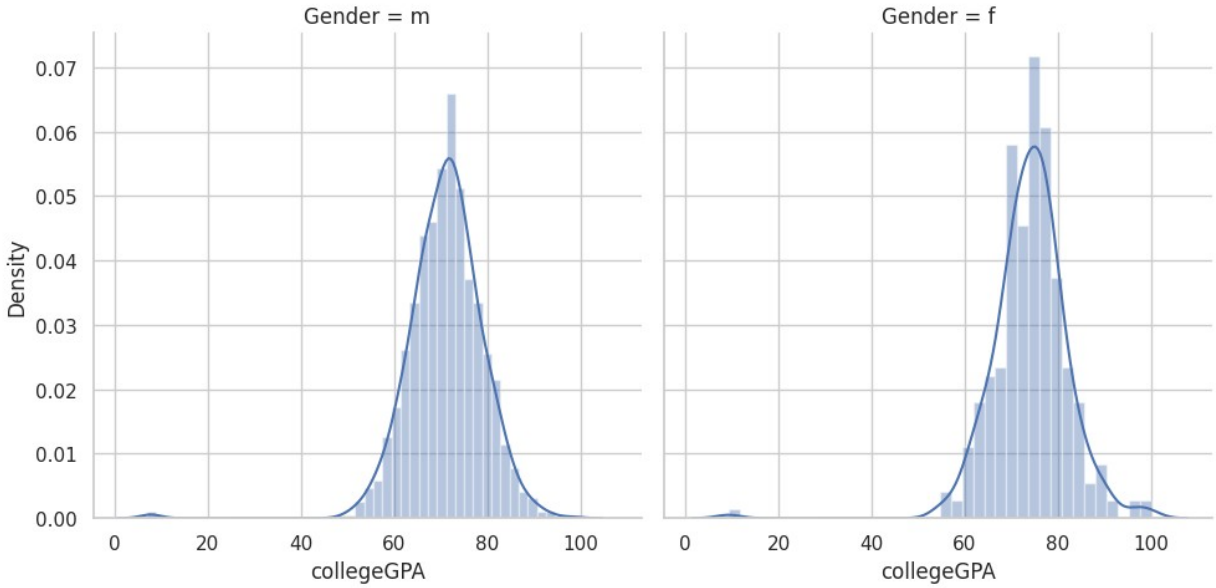



```
### What is average experience of software engineer and software
developer?
dataset[dataset.Designation.isin(['software engineer','software
developer']) & dataset.Gender=='m']['Period'].mean()
```

```
nan
```

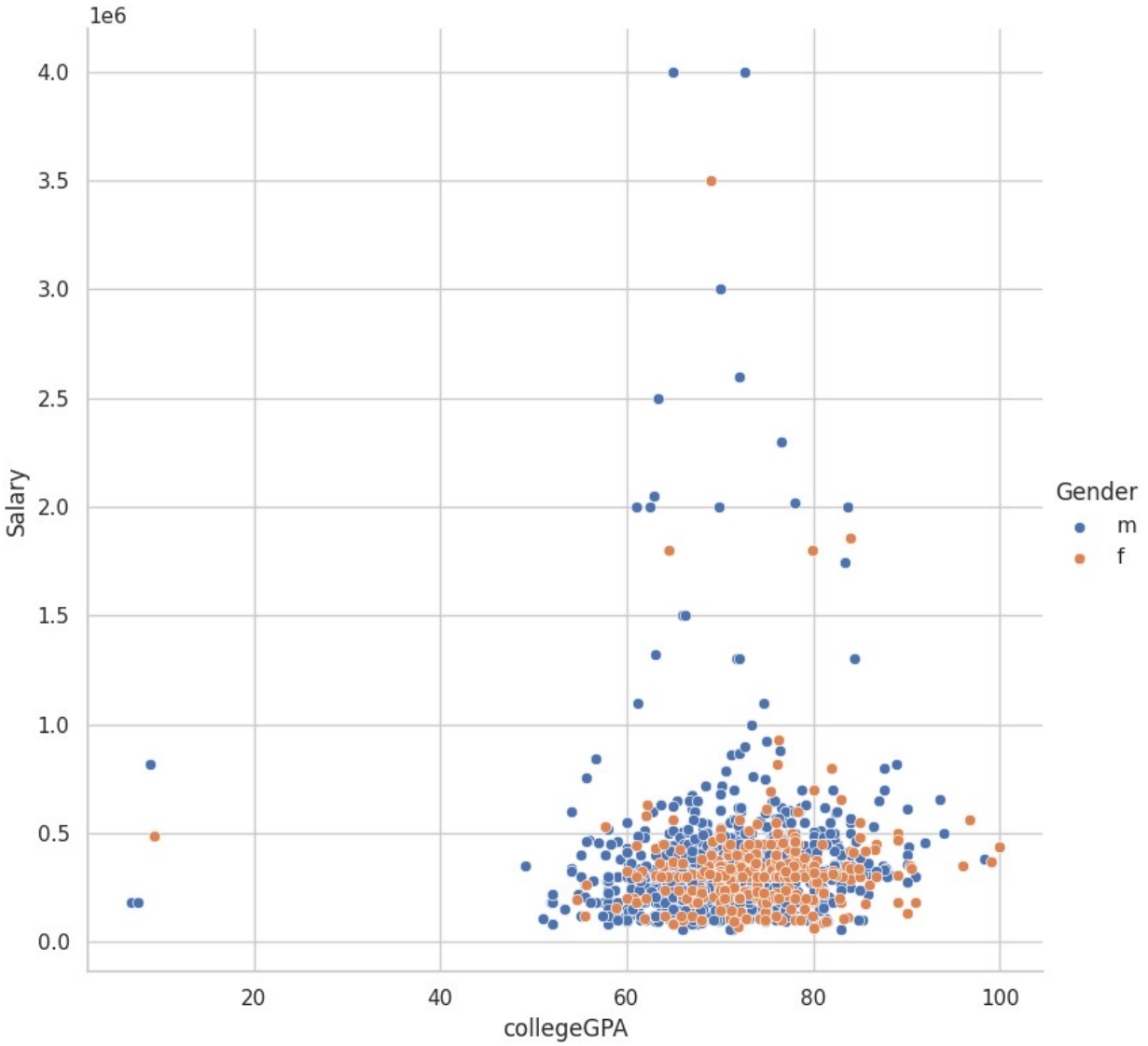
```
plt.figure(figsize=(10, 5))

# Creating a FacetGrid to visualize the distribution of college GPAs
for both genders
sns.FacetGrid(data=high_pay, col='Gender', height=5) \
    .map(sns.distplot, 'collegeGPA') \
    .add_legend()

plt.show()
```

```
<Figure size 1000x500 with 0 Axes>
```



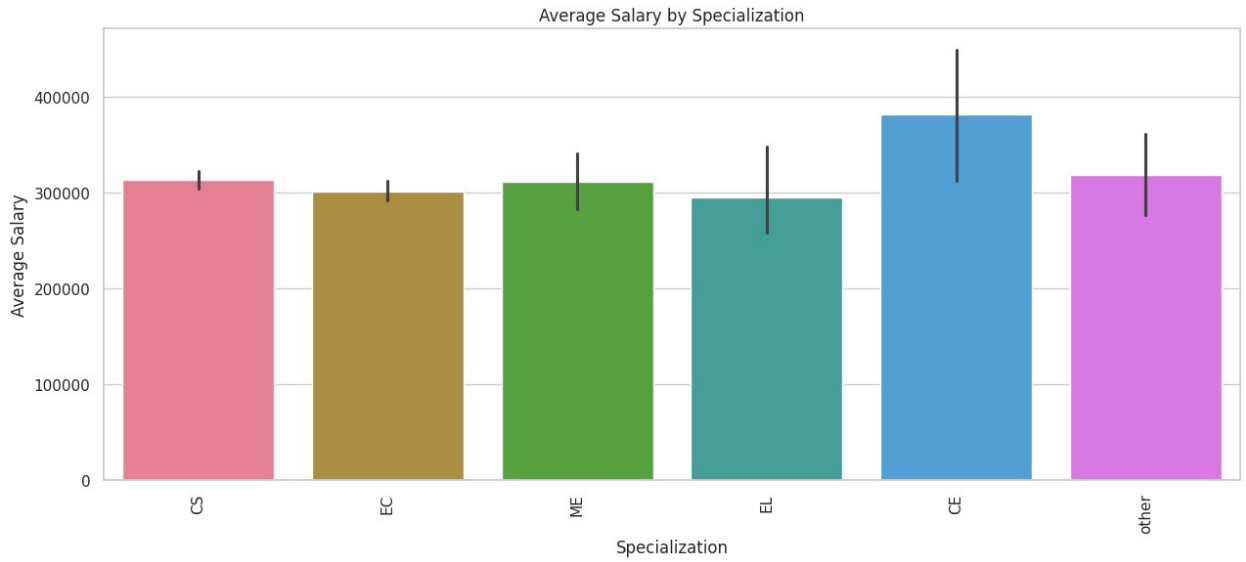

```
plt.figure(figsize=(10, 8))  # Adjust the figure size as needed

# Creating a FacetGrid to visualize the scatter plot of college GPA
vs. salary for both genders
sns.FacetGrid(data=high_pay, hue='Gender', height=8) \
    .map(sns.scatterplot, 'collegeGPA', 'Salary') \
    .add_legend()

plt.show()
```

```
<Figure size 1000x800 with 0 Axes>
```

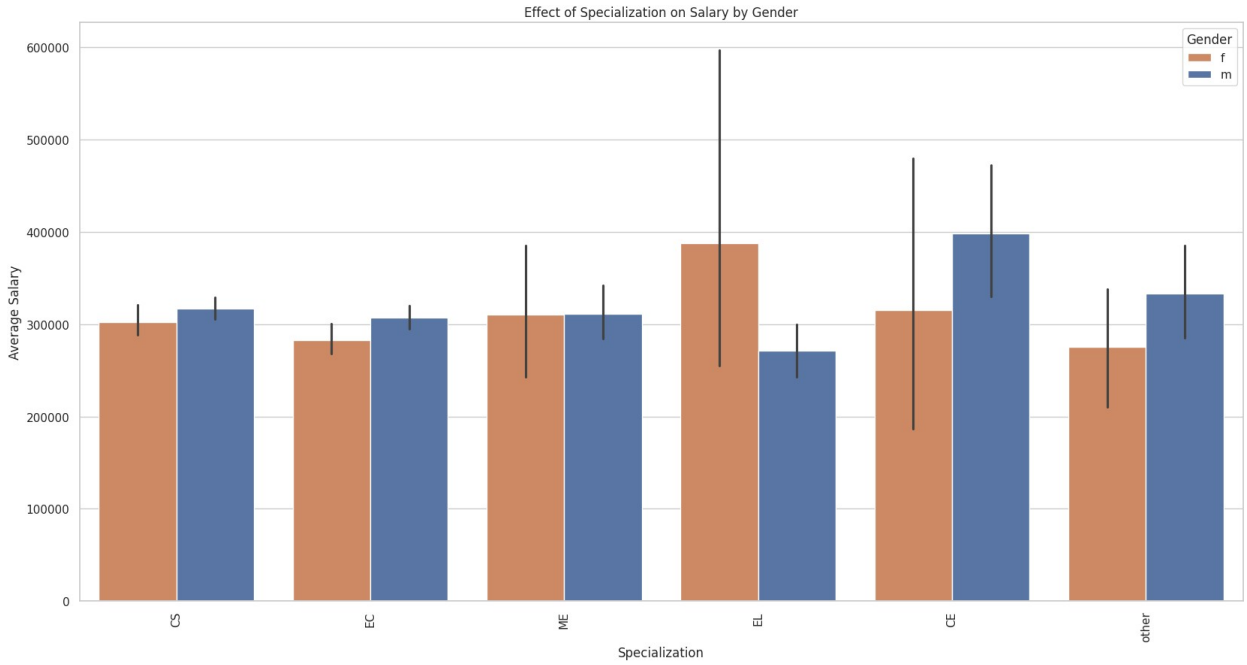



```
plt.figure(figsize=(15, 6))  # Adjust the figure size as needed

# Define a color palette with different colors for each specialization
palette = sns.color_palette("husl",
n_colors=len(dataset['Specialization'].unique()))

# Creating a bar plot of average salary by specialization with
specified palette
sns.barplot(data=dataset, x='Specialization', y='Salary',
palette=palette)

# Rotating x-axis labels for better readability
plt.xticks(rotation=90)

# Adding titles and labels
plt.title('Average Salary by Specialization')
```

```
plt.xlabel('Specialization')
plt.ylabel('Average Salary')

plt.show()
```

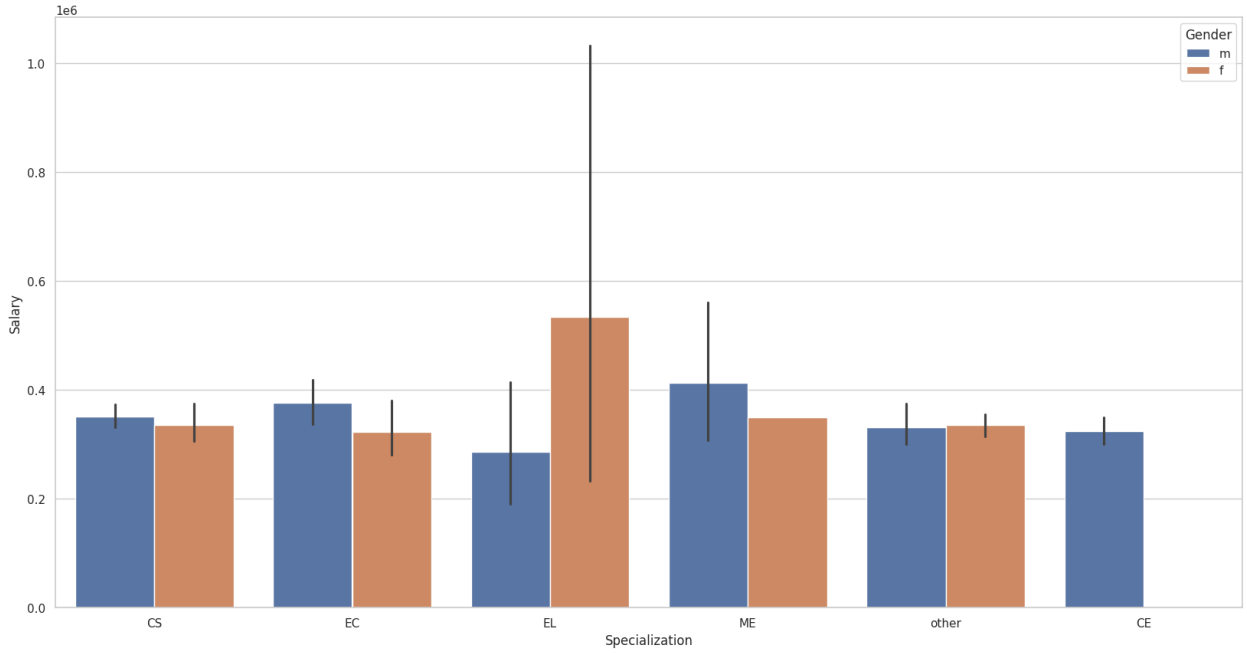



```
plt.figure(figsize=(20, 10))

# Define a custom palette with two colors for gender
palette = [(0.8666666666666667, 0.5176470588235295,
0.3215686274509804),
           (0.2980392156862745, 0.4470588235294118,
0.6901960784313725)]

# Creating a bar plot of average salary by specialization, with hue as
gender and custom palette
sns.barplot(data=dataset, x='Specialization', y='Salary',
hue='Gender', palette=palette)

# Adding titles and labels
plt.title('Effect of Specialization on Salary by Gender')
plt.xlabel('Specialization')
plt.ylabel('Average Salary')

plt.xticks(rotation=90)  # Rotating x-axis labels for better
readability
plt.legend(title='Gender')  # Adding a legend with title

plt.show()
```



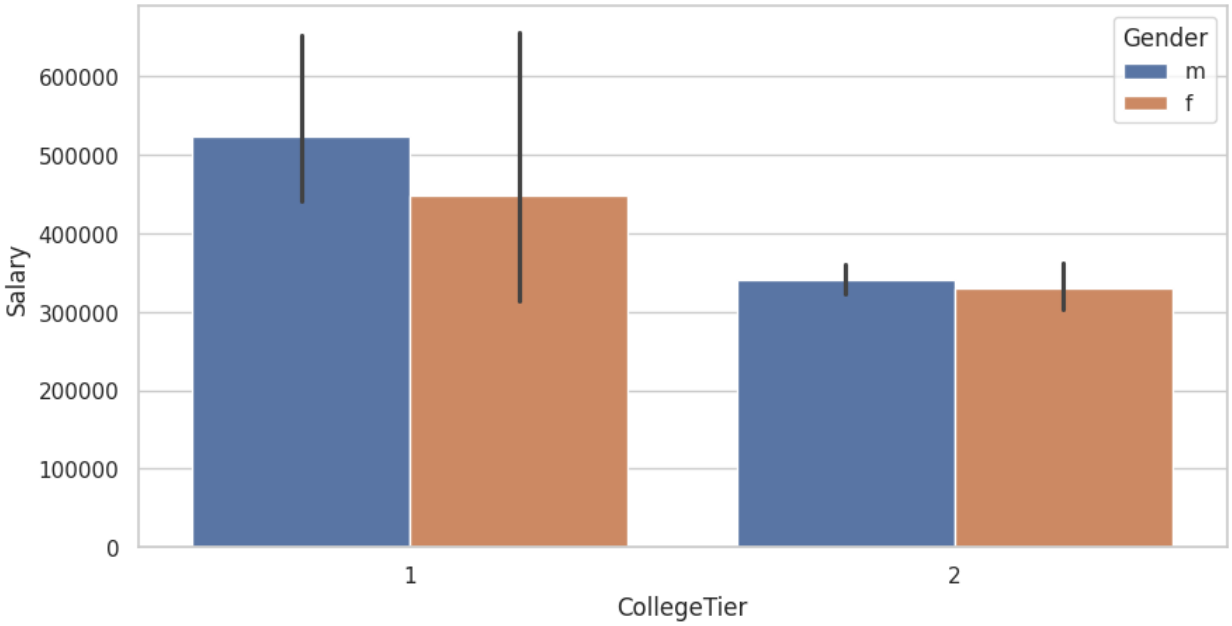

```
# for the dataset containing Highpaying Jobs
plt.figure(figsize=(20,10))
sns.barplot(data=high_pay,x='Specialization',y='Salary',hue='Gender')
```

```
<Axes: xlabel='Specialization', ylabel='Salary'>
```




```
### Lets us check salary with the College Tier
plt.figure(figsize=(10,5))
sns.barplot(data=high_pay,x='CollegeTier',y='Salary',hue='Gender')
```

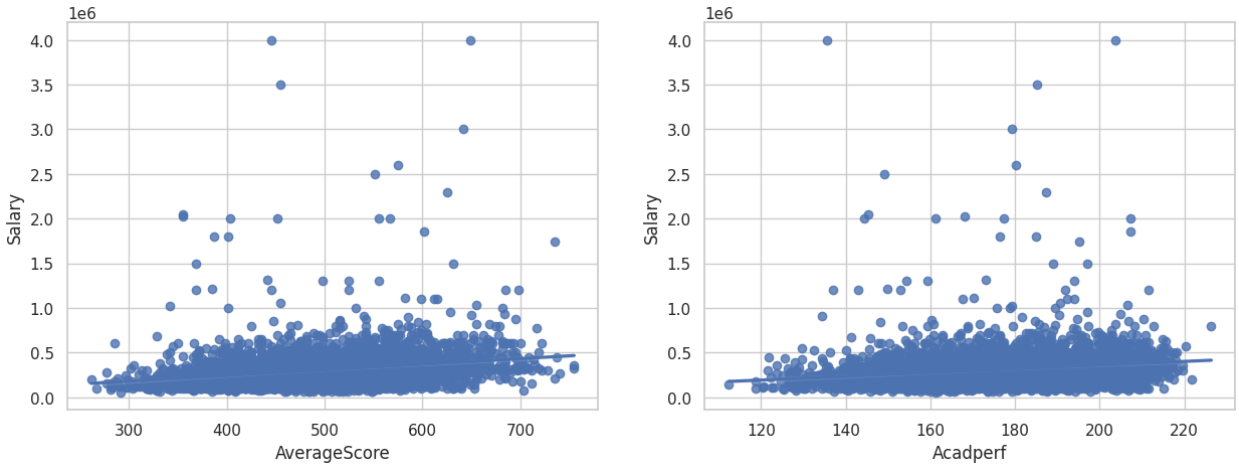
```
<Axes: xlabel='CollegeTier', ylabel='Salary'>
```




```
high_pay.groupby('CollegeTier').Gender.value_counts()
```

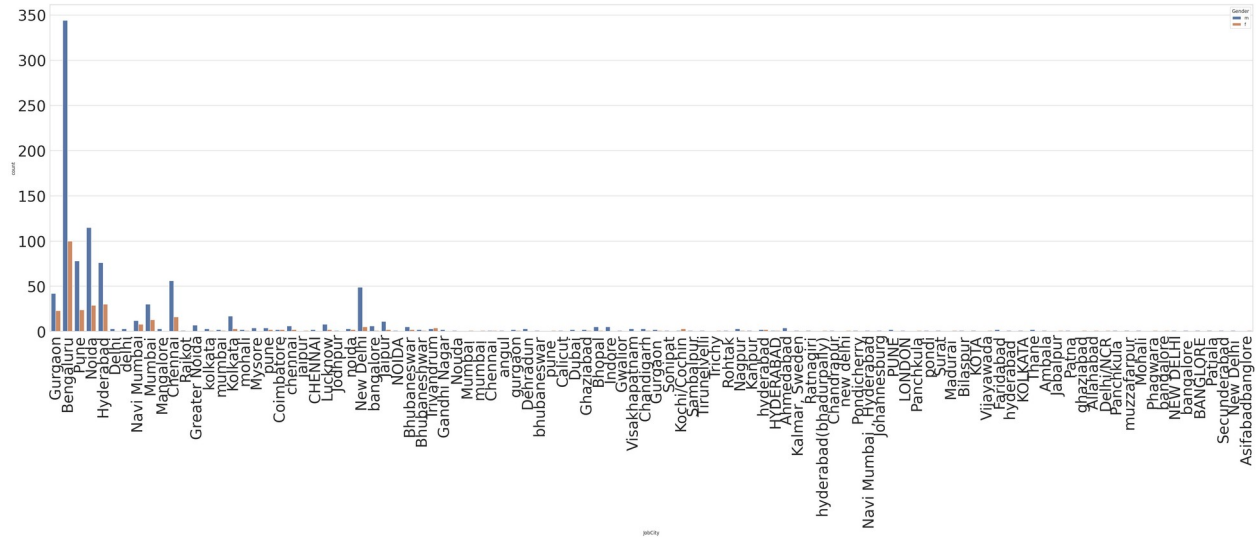
```
CollegeTier  Gender
1            m          85
             f          17
2            m         897
             f         287
Name: Gender, dtype: int64
```

```
plt.figure(figsize=(15,5))
dataset['AverageScore']=(dataset['Logical']+dataset['Quant']
+dataset['English'])/3
dataset['Acadperf']=dataset['10percentage']+dataset['12percentage']
+dataset['collegeGPA']/3
plt.subplot(1,2,1)
sns.regplot(x='AverageScore',y='Salary',data=dataset)
plt.subplot(1,2,2)
sns.regplot(x='Acadperf',y='Salary',data=dataset)
plt.show()
```



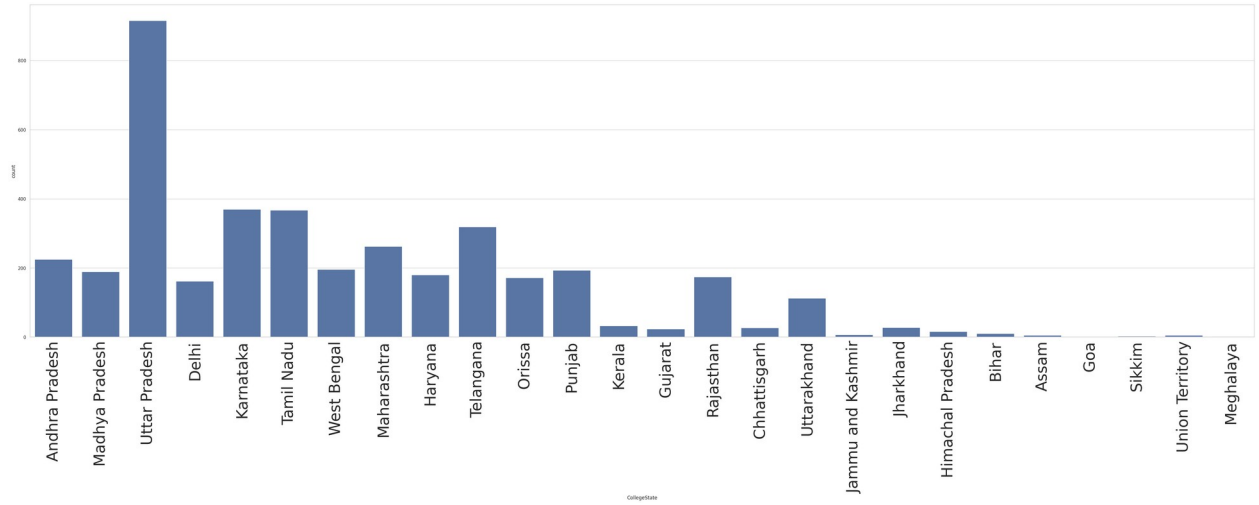

```
plt.figure(figsize=(55,15))
sns.countplot(x="JobCity",data=high_pay,hue="Gender")
plt.xticks(fontsize=38,rotation=90)
plt.yticks(fontsize=38)
```

```
(array([  0.,  50., 100., 150., 200., 250., 300., 350., 400.]),
 [Text(0, 0.0, '0'),
  Text(0, 50.0, '50'),
  Text(0, 100.0, '100'),
  Text(0, 150.0, '150'),
  Text(0, 200.0, '200'),
  Text(0, 250.0, '250'),
  Text(0, 300.0, '300'),
  Text(0, 350.0, '350'),
  Text(0, 400.0, '400')])
```

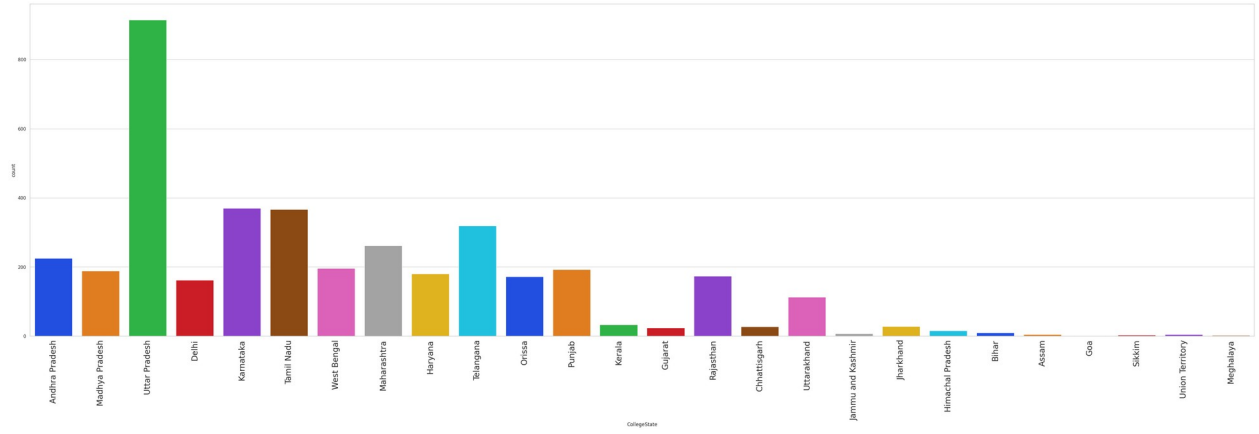



```
plt.figure(figsize=(55,15))
sns.countplot(data=dataset,x='CollegeState')
```

```
plt.xticks(fontsize=38,rotation=90)
plt.show()
```

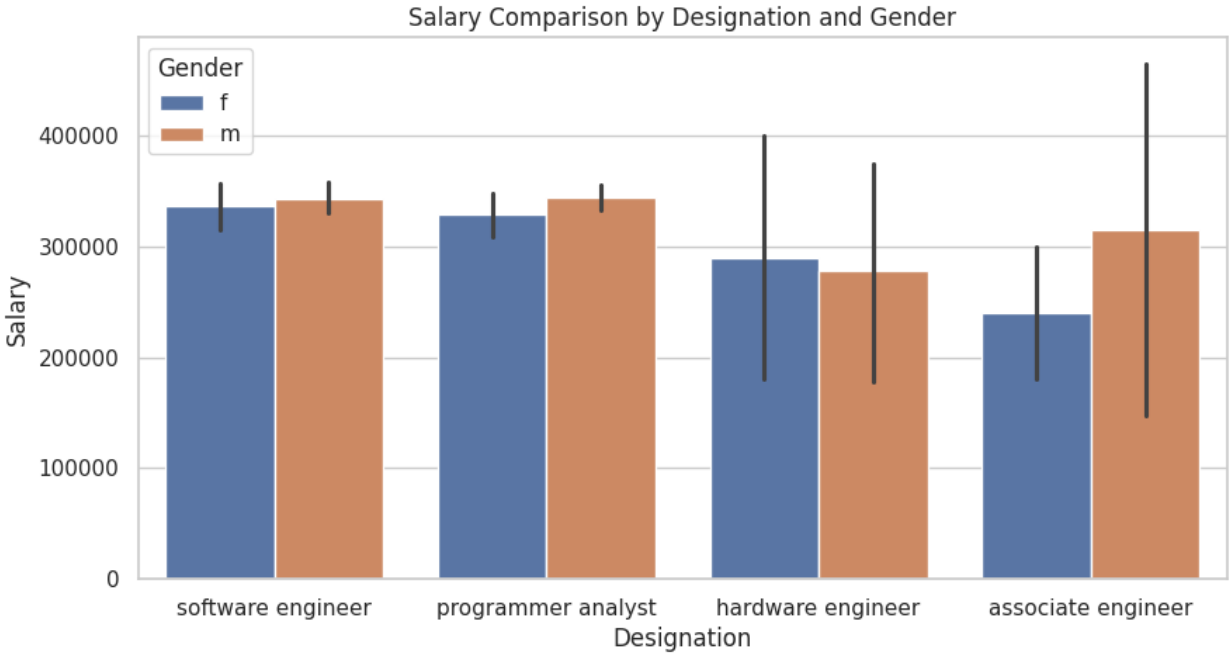



```
plt.figure(figsize=(55, 15))

# Creating a count plot of CollegeState with different colors for each state
sns.countplot(data=dataset, x='CollegeState', palette='bright')

# Adjusting the font size and rotation of x-axis labels
plt.xticks(fontsize=20, rotation=90)

plt.show()
```




```
# Filter the dataset based on specific designations
new = dataset[dataset["Designation"].isin(["programmer analyst",
"software engineer", "hardware engineer", "associate engineer"])]

plt.figure(figsize=(10, 5))
```

```
# Creating a bar plot comparing salaries among different job
designations while considering gender
sns.barplot(x="Designation", y="Salary", hue="Gender", data=new)

# Adding titles and labels
plt.title('Salary Comparison by Designation and Gender')
plt.xlabel('Designation')
plt.ylabel('Salary')

plt.show()
```

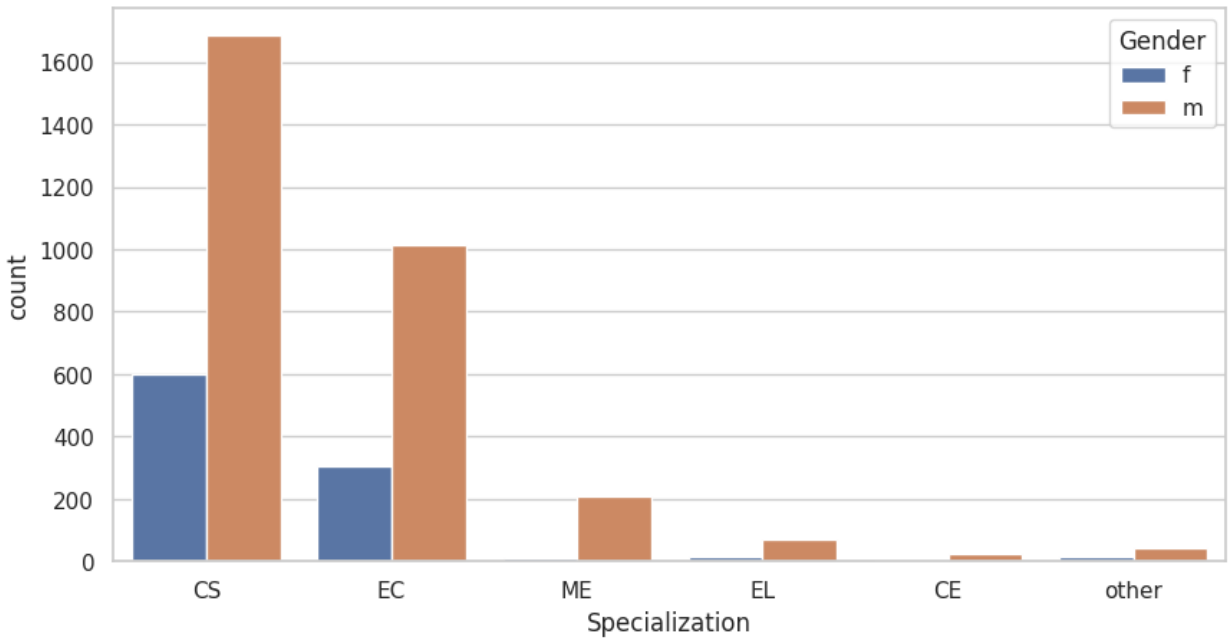



```
from scipy import stats as st
from scipy.stats import chi2_contingency as cst
pv = st.ttest_1samp(new['Salary'],popmean=250000)[1]
### for a 95% confidence interval,my p- value should be >0.05 to claim
the null hypothesis
if pv < 0.05:
 print('We reject the null hypothesis and Average salary is not equal
to 250k')
else:
 print('We fail to reject null hypothesis and Avergae salary is equal
to 250k')
```

```
We reject the null hypothesis and Average salary is not equal to 250k
```

```
plt.figure(figsize=(10,5))
sns.countplot(data=dataset,x='Specialization',hue='Gender')
```

```
<Axes: xlabel='Specialization', ylabel='count'>
```




```
sample_columns =
pd.crosstab(dataset['Gender'],dataset['Specialization'],margins=True)
pv = cst(sample_columns)[1]
if pv < 0.05:
 print('We reject the null hypothesis and Gender impacts
specialization')
else:
 print('We fail to reject null hypothesis and Gender does not impact
specialization')
```

```
We reject the null hypothesis and Gender impacts specialization
```