Manqiong

2017 State of Data Science - Kaggle survey

0 voters

forked from 2017 State of Data Science - Kaggle survey by Anisotropic (+94/−173)

last run a minute ago · Python notebook · 10 views

using data from multiple data sources · Public

Tags multiple data sources Add Tag

Notebook

In [1]:
```
import numpy as np
import pandas as pd
import scipy as sp
import csv
%matplotlib inline
import matplotlib.pylab as plt

import pandas as pd
import matplotlib.pyplot as plt; plt.rcdefaults()
import numpy as np
import matplotlib.pyplot as plt

import seaborn as sns
import matplotlib.ticker as ticker
```

In [2]:
```
mcr=pd.read_csv('../input/kaggle-survey-2017/multipleChoiceResponses.csv',
 encoding="ISO-8859-1", low_memory=False)
mcr.head()
```

Out[2]:

	GenderSelect	Country	Age	EmploymentStatus	StudentStatus	LearningDataScience	CodeWri
0	Non-binary, genderqueer, or gender non-conforming	NaN	NaN	Employed full-time	NaN	NaN	Yes
1	Female	United States	30.0	Not employed, but looking for work	NaN	NaN	NaN
2	Male	Canada	28.0	Not employed, but looking for work	NaN	NaN	NaN
3	Male	United States	56.0	Independent contractor, freelancer, or self-em...	NaN	NaN	Yes
4	Male	Taiwan	38.0	Employed full-time	NaN	NaN	Yes

5 rows × 228 columns

In [3]:
```
cr=pd.read_csv('../input/kaggle-survey-2017/conversionRates.csv', encoding
="ISO-8859-1", low_memory=False)
cr.head()
```

Out[3]:

	Unnamed: 0	originCountry	exchangeRate
0	1	USD	1.000000
1	2	EUR	1.195826
2	3	INR	0.015620
3	4	GBP	1.324188
4	5	BRL	0.321350

In [4]:
```
fr=pd.read_csv('../input/kaggle-survey-2017/freeformResponses.csv', encodi
ng="ISO-8859-1", low_memory=False)
fr.head()
```

Out[4]:

	GenderFreeForm	KaggleMotivationFreeForm	CurrentJobTitleFreeForm	MLToolNextYearFreeForm	M
0	NaN	NaN	NaN	NaN	N
1	NaN	NaN	NaN	NaN	N
2	NaN	NaN	teacher	NaN	N
3	NaN	NaN	NaN	NaN	N
4	NaN	NaN	NaN	NaN	N

5 rows × 62 columns

In [5]:
```
sch=pd.read_csv('../input/kaggle-survey-2017/schema.csv', encoding="ISO-88
59-1", low_memory=False)
sch.head()
```

Out[5]:

	Column	Question	Asked
0	GenderSelect	Select your gender identity. - Selected Choice	All

1	GenderFreeForm	Select your gender identity. - A different ide...	All
2	Country	Select the country you currently live in.	All
3	Age	What's your age?	All
4	EmploymentStatus	What's your current employment status?	All

In [6]:
```
sch.tail()
```

Out[6]:

	Column	Question	Asked
285	JobFactorRemote	How are you assessing potential job opportunit...	Learners
286	JobFactorIndustry	How are you assessing potential job opportunit...	Learners
287	JobFactorLeaderReputation	How are you assessing potential job opportunit...	Learners
288	JobFactorDiversity	How are you assessing potential job opportunit...	Learners
289	JobFactorPublishingOpportunity	How are you assessing potential job opportunit...	Learners

In [7]:
```
sch.Asked.value_counts()
```

Out[7]:
```
CodingWorker       161
All                 70
Learners            41
Worker1              6
CodingWorker-NC      5
OnlineLearners       2
Worker               2
Non-worker           2
Non-switcher         1
Name: Asked, dtype: int64
```

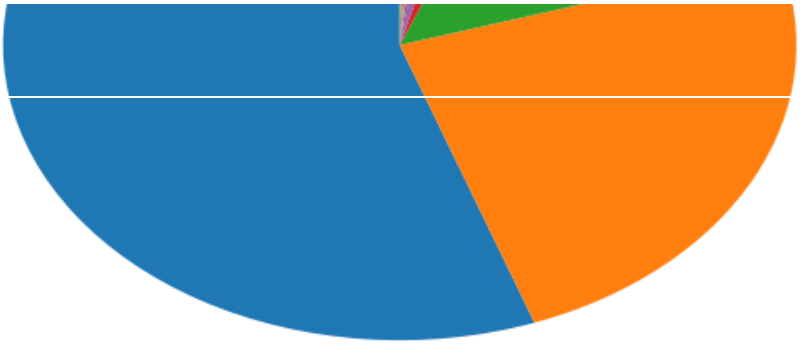
In [8]:
```
sch.groupby('Asked').size()
```

Out[8]:
```
Asked
All                  70
CodingWorker        161
```

```
CodingWorker        161
CodingWorker-NC       5
Learners             41
Non-switcher          1
Non-worker            2
OnlineLearners        2
Worker                2
Worker1               6
dtype: int64
```

In [9]:

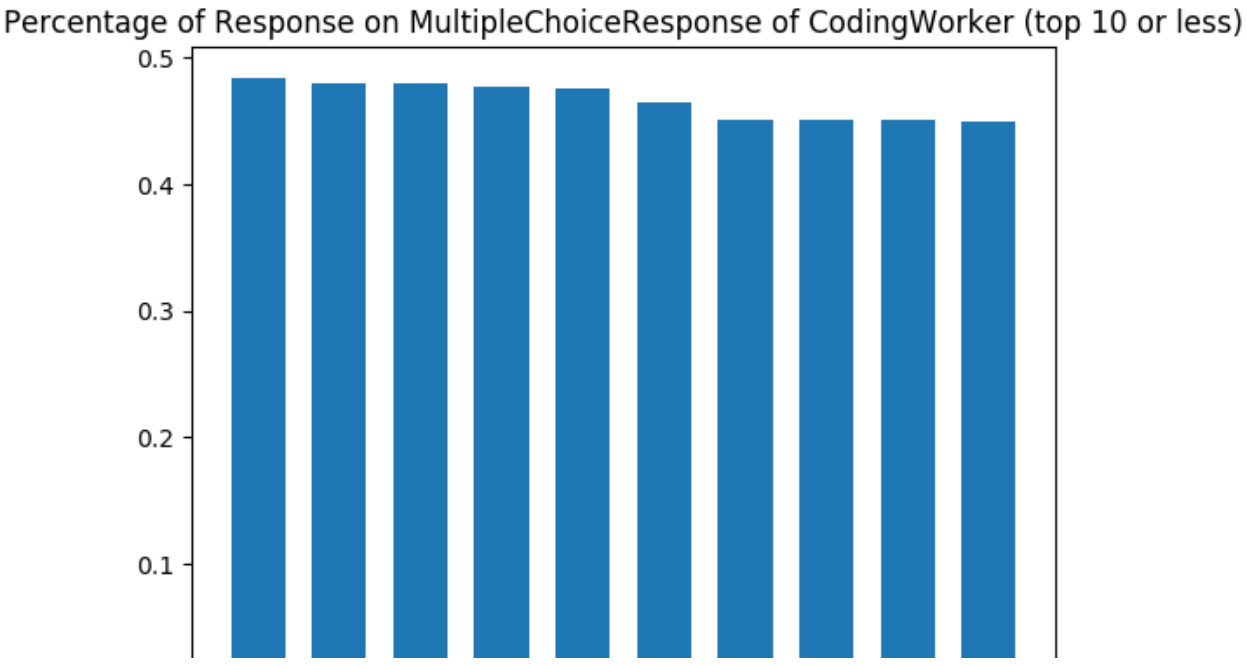
```
#determine the number of categories in Aksed
Piedata=sch.Asked.value_counts()

labels=[]
for i in Piedata.index:
    labels.append('{0}'.format(i))
labels

x = labels
y = Piedata[:]

percent = 100.*y/y.sum()

patches, texts = plt.pie(y, startangle=90, radius=1.2)
labels = ['{0} - {1:1.2f} %'.format(i,j) for i,j in zip(x, percent)]

sort_legend = True
if sort_legend:
    patches, labels, dummy =  zip(*sorted(zip(patches, labels, y),
                                          key=lambda x: x[2],
                                          reverse=True))

plt.legend(patches, labels, loc='best', bbox_to_anchor=(-0.1, 1.),
           fontsize=8)
plt.show()
```

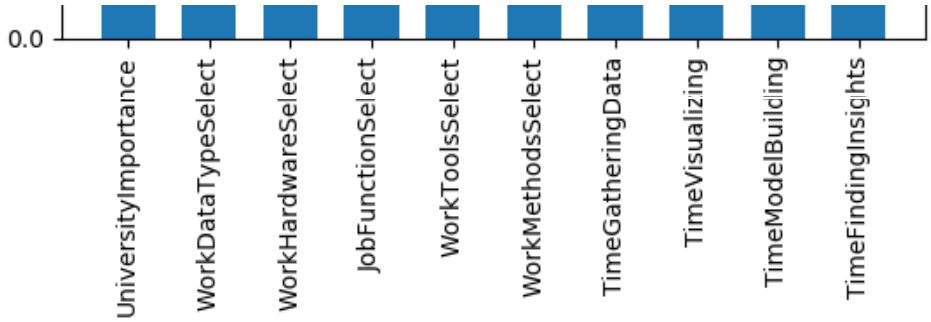


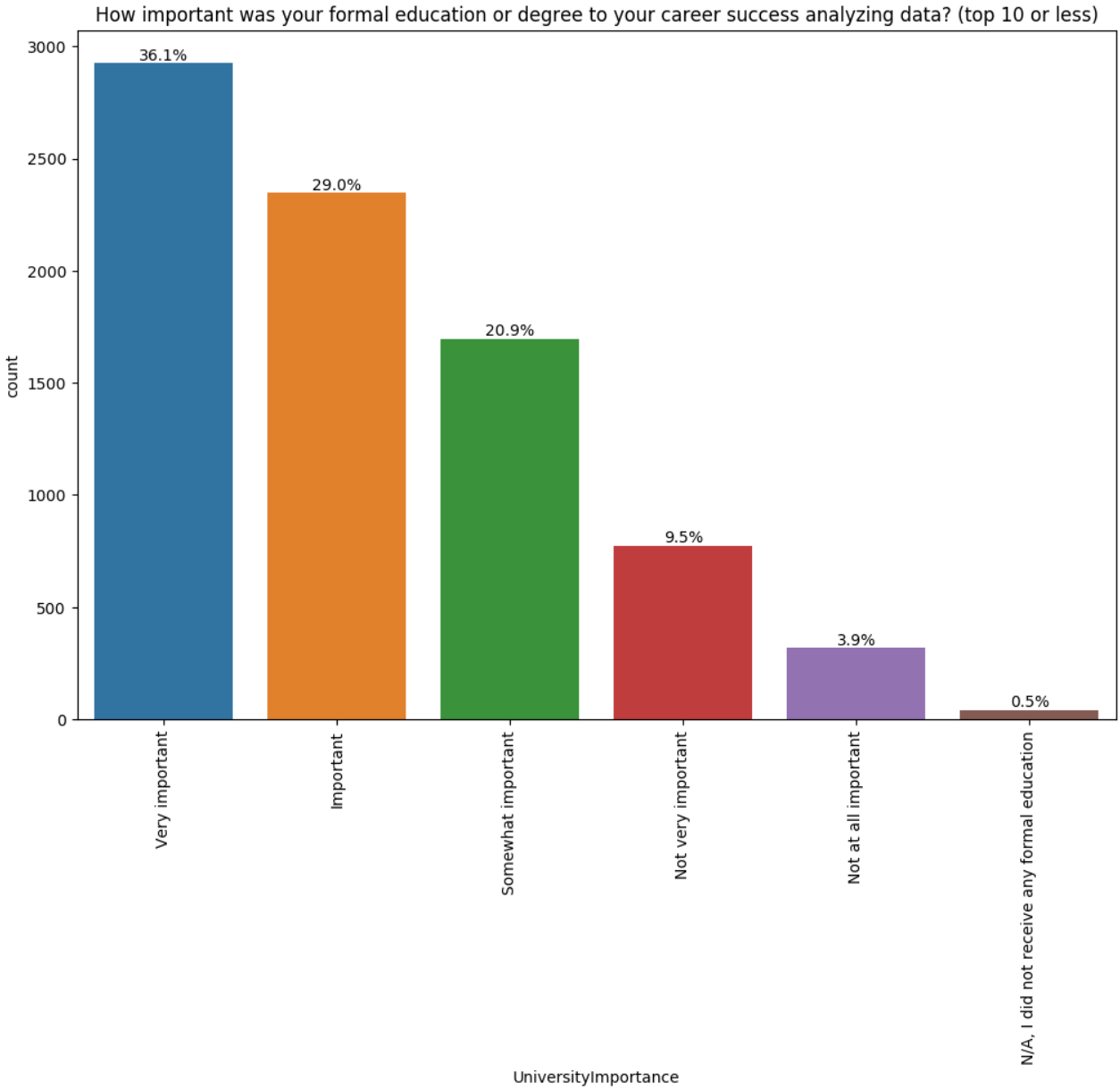
In [10]:
```
# Group by Asked on MultipleChoiceResponse and FreeFormResponse
Asked_N=range(len(sch.Asked.value_counts().index))
Askgp_mcr=[[] for z in Asked_N]
Askgp_fr=[[] for z in Asked_N]
for i in Asked_N:
    Askgp=list(sch['Column'][sch['Asked']==sch.Asked.value_counts().index[
i]])
    for j in Askgp:
        if j not in fr:
            Askgp_mcr[i].append(j)
        else:
            Askgp_fr[i].append(j)
```

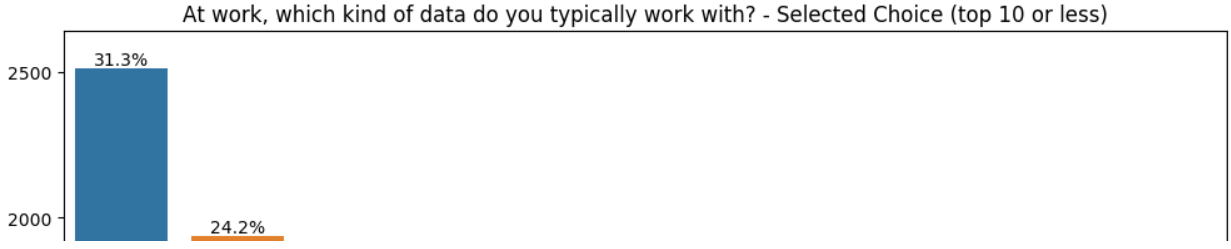
In [11]:
```
mcr_perc_resp=[]


for i in range(len(sch.Asked.value_counts().index)):

    # Percentage of Response of each questions on MultipleChoiceResponse
    perc_resp=mcr[Askgp_mcr[i]].count()/(mcr[Askgp_mcr[i]].count()+mcr[Ask
gp_mcr[i]].isnull().sum())
    perc_resp_sort=perc_resp.sort_values(ascending=False)
    df_perc_resp=perc_resp_sort.to_frame() # You need to create a datafram
e first, then you can append to a list.
    mcr_perc_resp.append(df_perc_resp)

    y = mcr_perc_resp[i].iloc[:10][0]
    N = len(y)
    x = range(N)
    width = 1/1.5
```

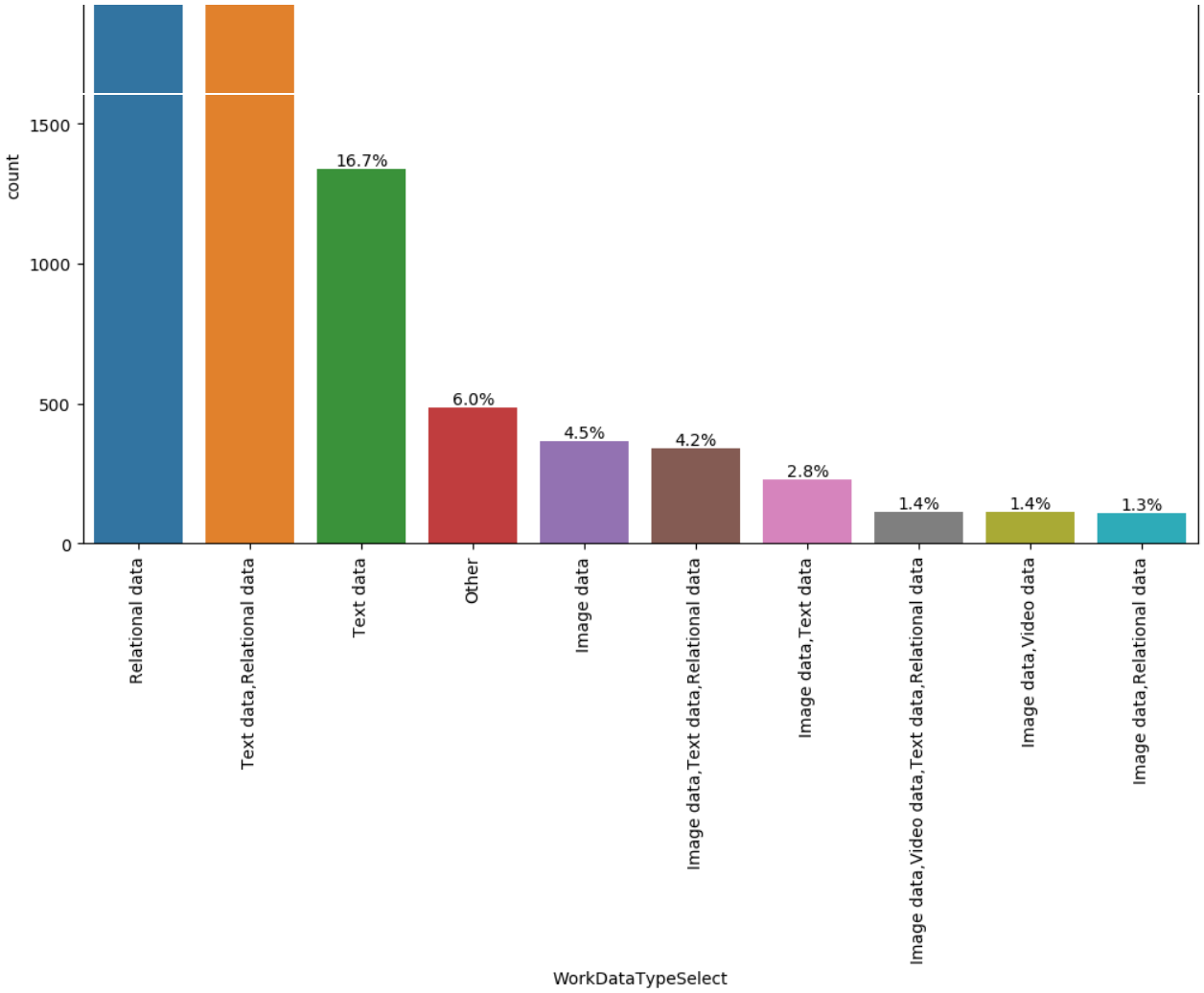
```
    plt.title('Percentage of Response on MultipleChoiceResponse of {0} (to
p 10 or less)'.format(sch.Asked.value_counts().index[i]))
    plt.xticks(x, list(mcr_perc_resp[i].index)[:10],rotation='vertical')
    ax=plt.bar(x, y, width)

    plt.show()

    for j in range(len(y)):
        plt.figure(figsize=(12,8))
        ax2 = sns.countplot(x=y.index[j], data=mcr, order=mcr[y.index[j]].
value_counts().index[:10])
        plt.title('{0} (top 10 or less)'.format(list(sch['Question'][sch[
'Column']==y.index[j]])[0]))
        plt.xlabel(y.index[j])
        plt.xticks(rotation=90)

        ncount = len(mcr[y.index[j]].dropna())

        for p in ax2.patches:
            x2=p.get_bbox().get_points()[:,0]
            y2=p.get_bbox().get_points()[1,1]
            ax2.annotate('{:.1f}%'.format(100.*y2/ncount), (x2.mean(), y2
),
                    ha='center', va='bottom') # set the alignment of the t
ext


        plt.show()
```

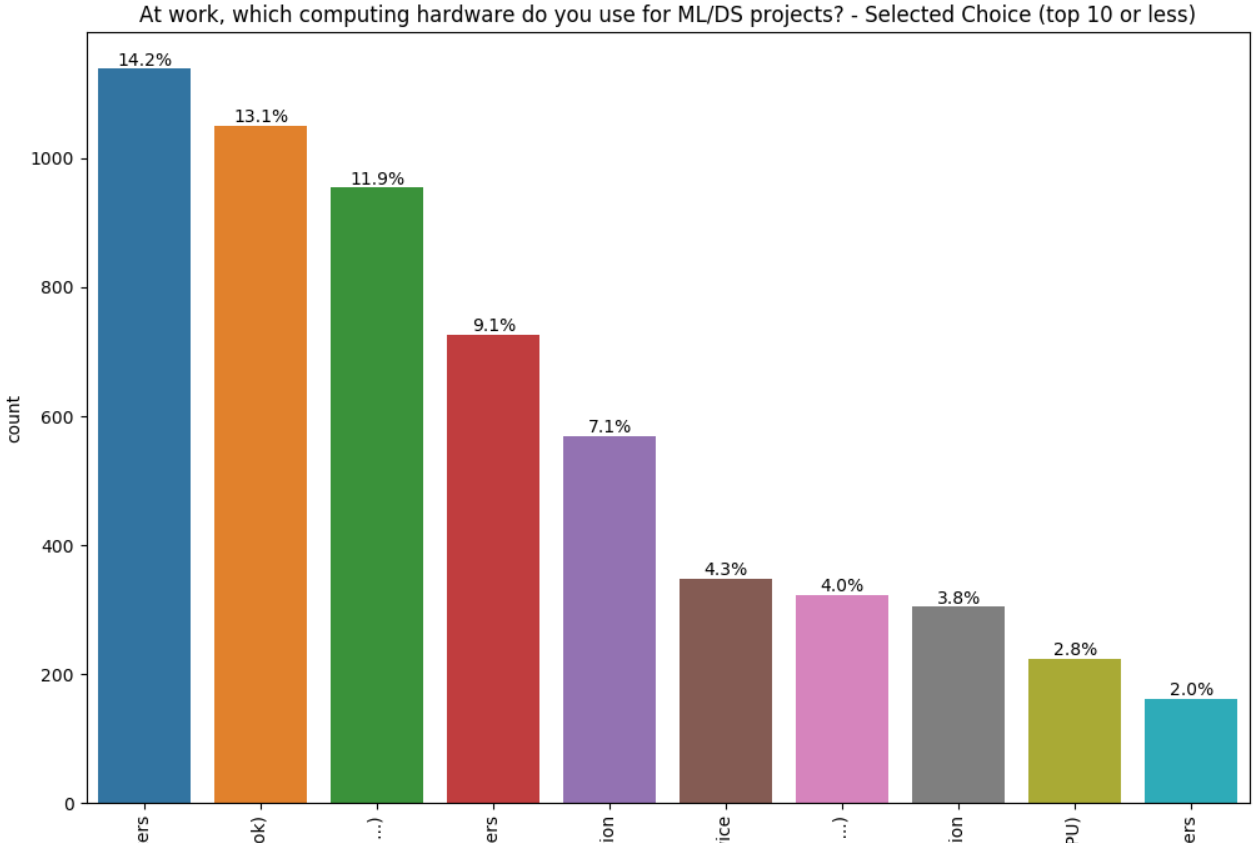


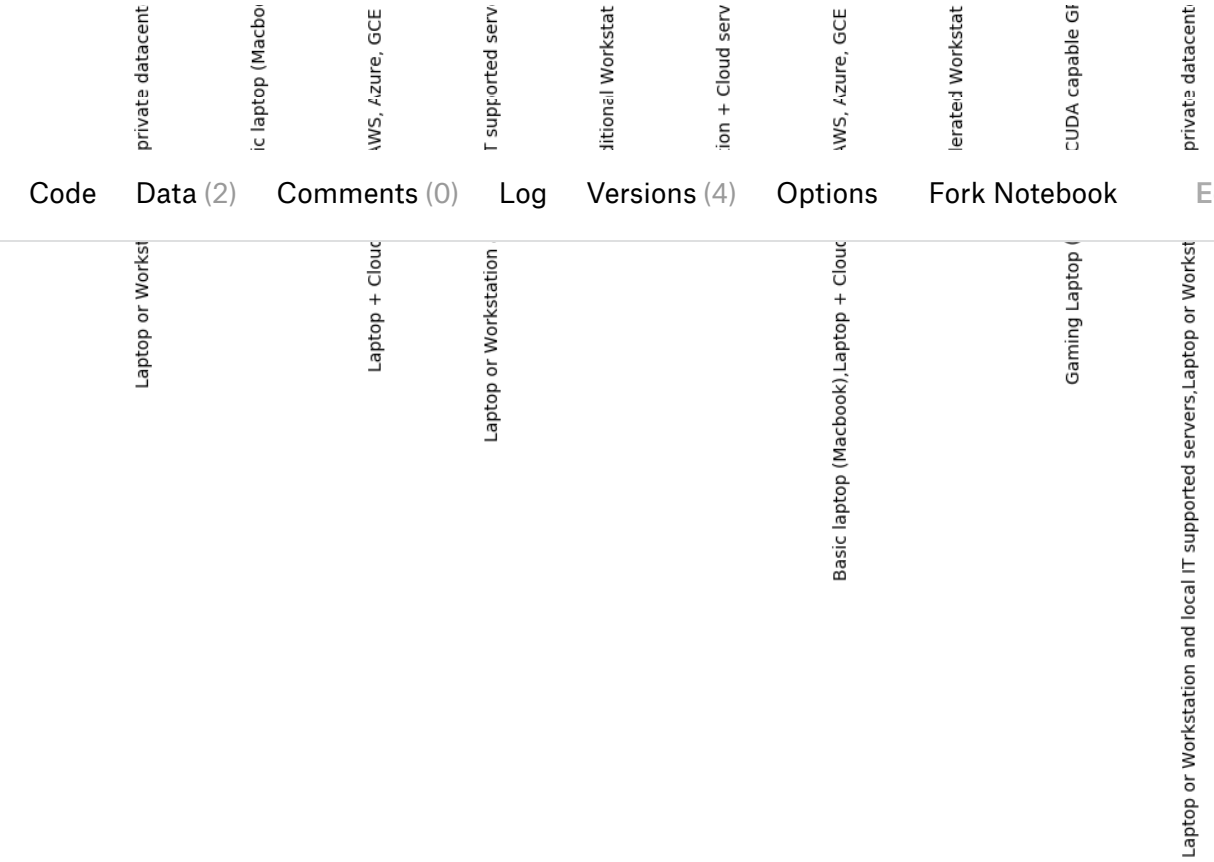

0.0
UniversityImportance
WorkDataTypeSelect
WorkHardwareSelect
JobFunctionSelect
WorkToolsSelect
WorkMethodsSelect
TimeGatheringData
TimeVisualizing
TimeModelBuilding
TimeFindingInsights

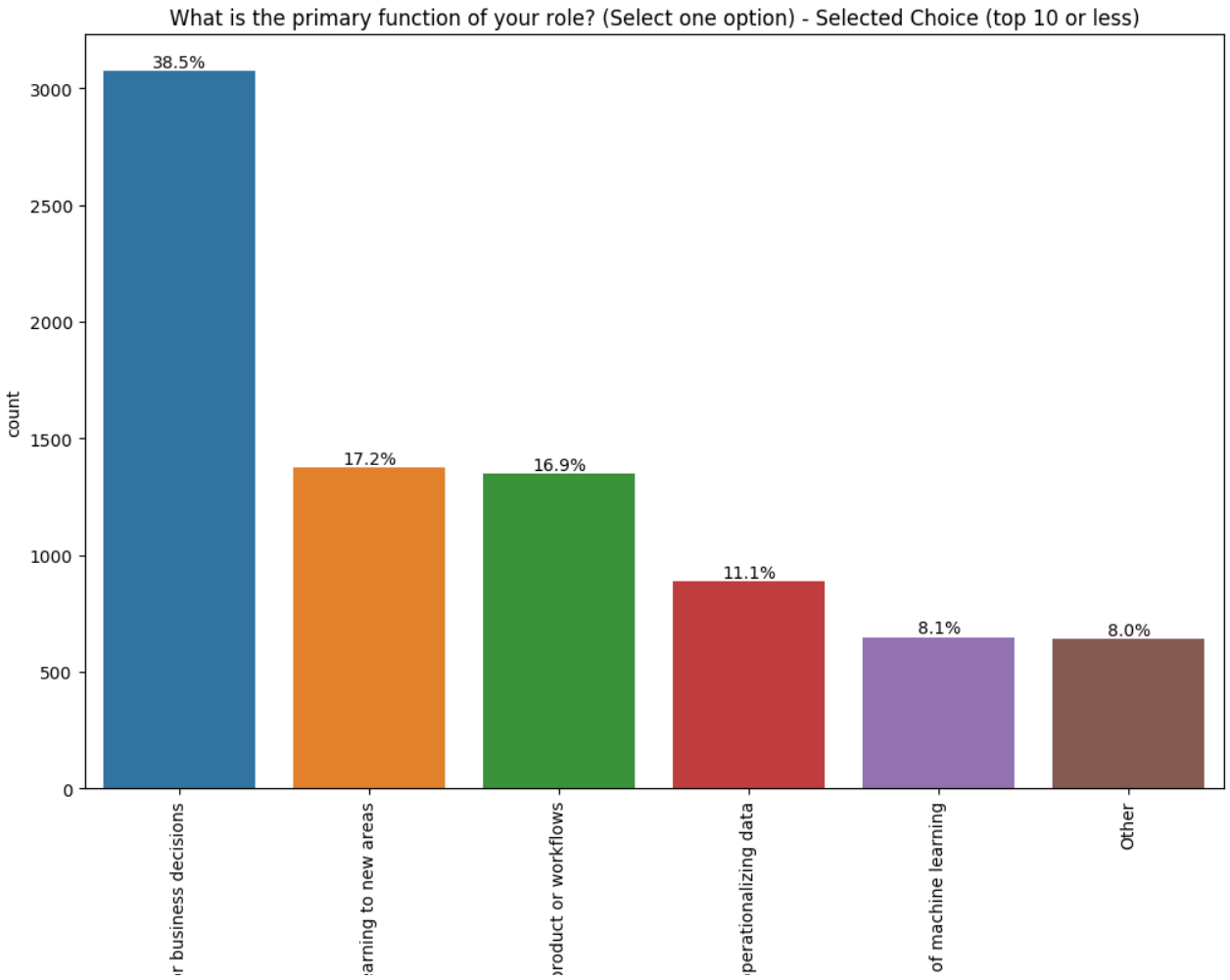

How important was your formal education or degree to your career success analyzing data? (top 10 or less)
count
3000
2500
2000
1500
1000
500
0
36.1%
29.0%
20.9%
9.5%
3.9%
0.5%
Very important
Important
Somewhat important
Not very important
Not at all important
N/A, I did not receive any formal education
UniversityImportance

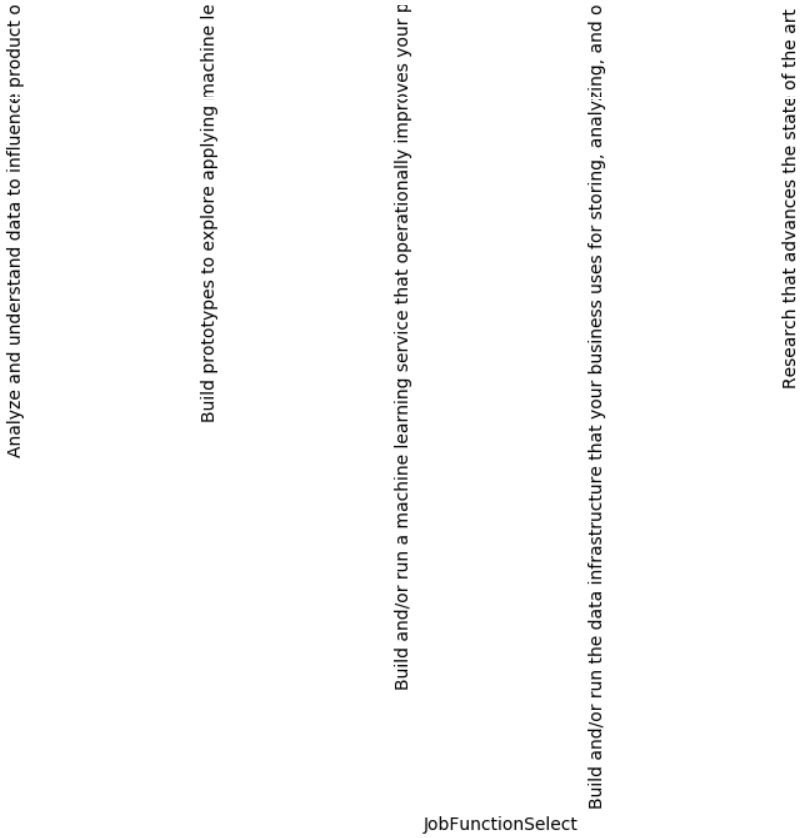

At work, which kind of data do you typically work with? - Selected Choice (top 10 or less)
2500
2000
31.3%
24.2%


count
1500
1000
500
0
16.7%
6.0%
4.5%
4.2%
2.8%
1.4%
1.4%
1.3%
Relational data
Text data,Relational data
Text data
Other
Image data
Image data,Text data,Relational data
Image data,Text data
Image data,Video data,Text data,Relational data
Image data,Video data
Image data,Relational data
WorkDataTypeSelect

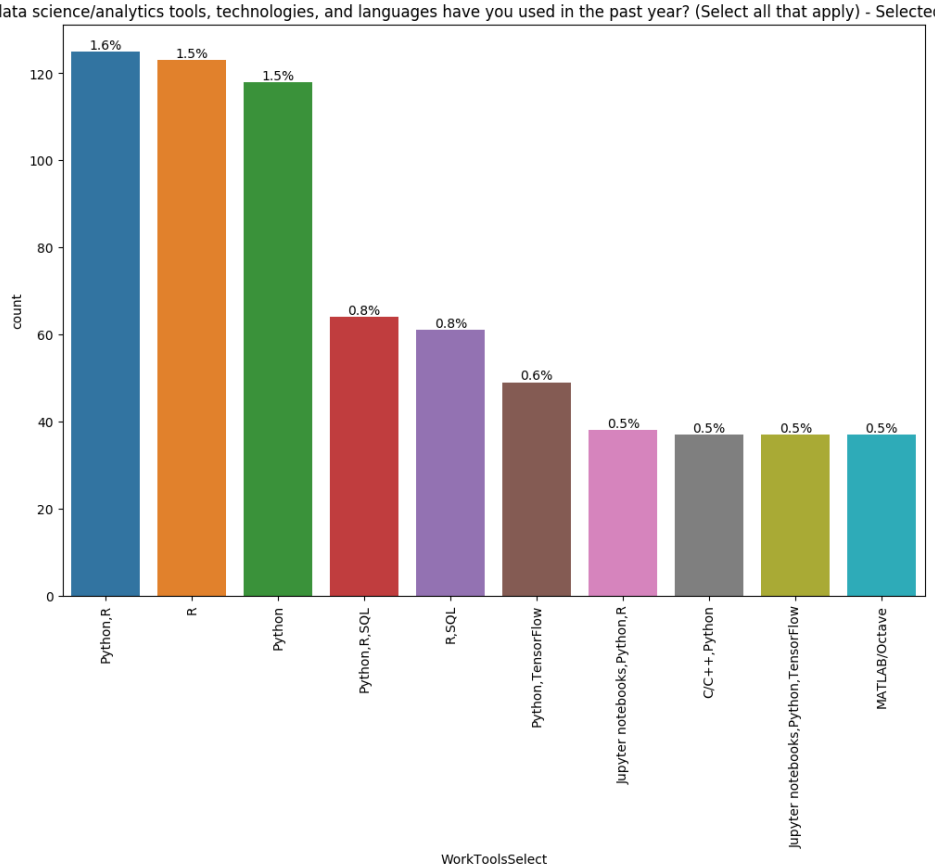

At work, which computing hardware do you use for ML/DS projects? - Selected Choice (top 10 or less)
count
1000
800
600
400
200
0
14.2%
13.1%
11.9%
9.1%
7.1%
4.3%
4.0%
3.8%
2.8%
2.0%
ers
ok)
...)
ers
ion
ice
...)
ion
PU)
ers


private datacent
ic laptop (Macbo
WS, Azure, GCE
T supported serv
ditional Workstat
ion + Cloud serv
WS, Azure, GCE
lerated Workstat
CUDA capable GI
private datacent
Laptop or Workst
Laptop + Cloud
Laptop or Workstation
Basic laptop (Macbook),Laptop + Cloud
Gaming Laptop
Laptop or Workstation and local IT supported servers,Laptop or Workst
WorkHardwareSelect

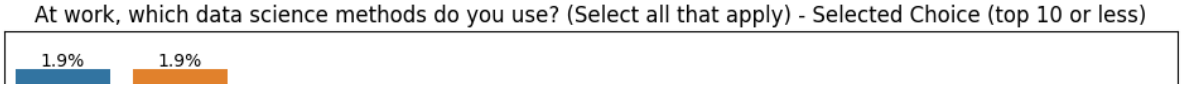

What is the primary function of your role? (Select one option) - Selected Choice (top 10 or less)
count
0
500
1000
1500
2000
2500
3000
38.5%
17.2%
16.9%
11.1%
8.1%
8.0%
r business decisions
arning to new areas
oduct or workflows
perationalizing data
of machine learning
Other

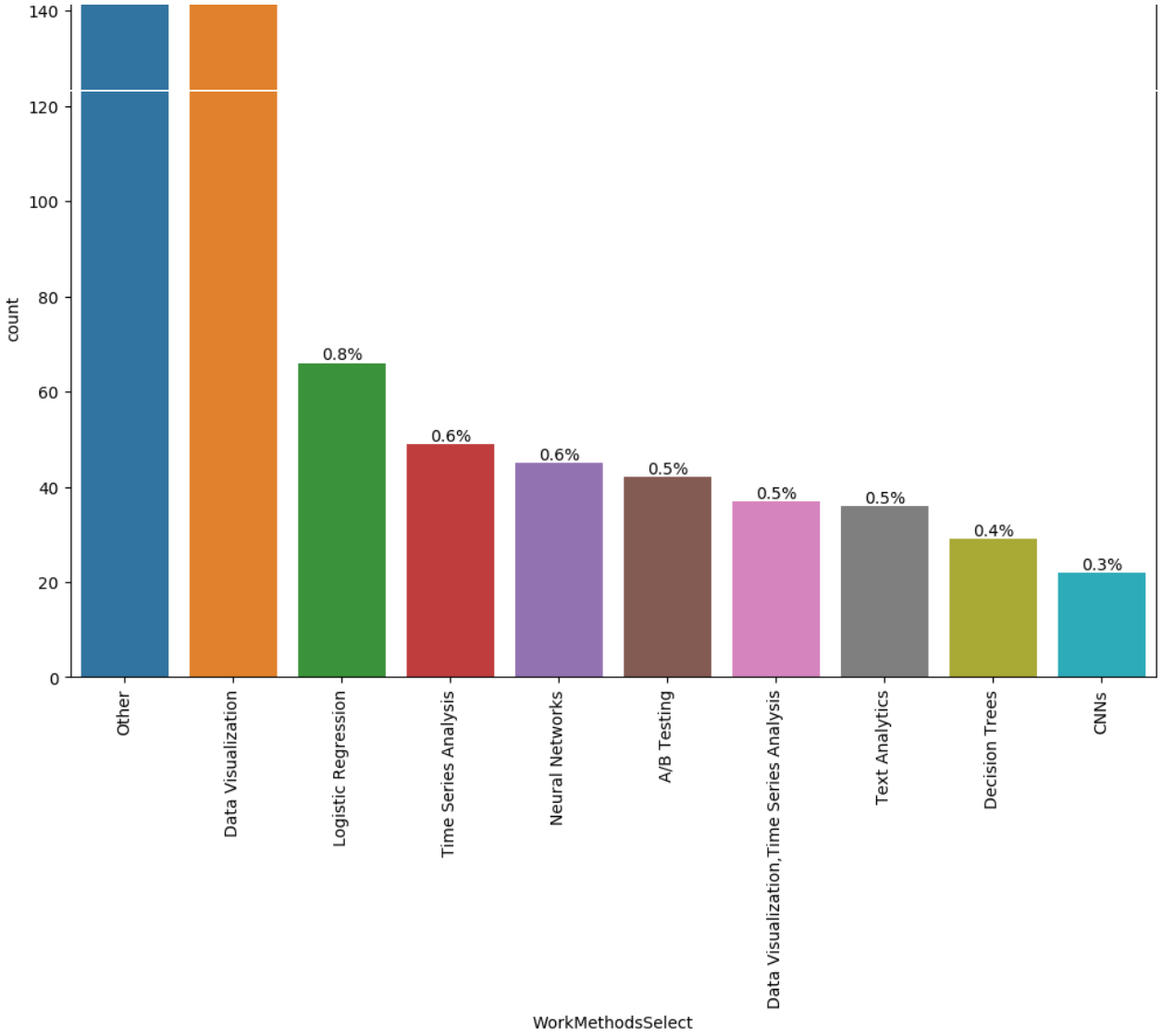


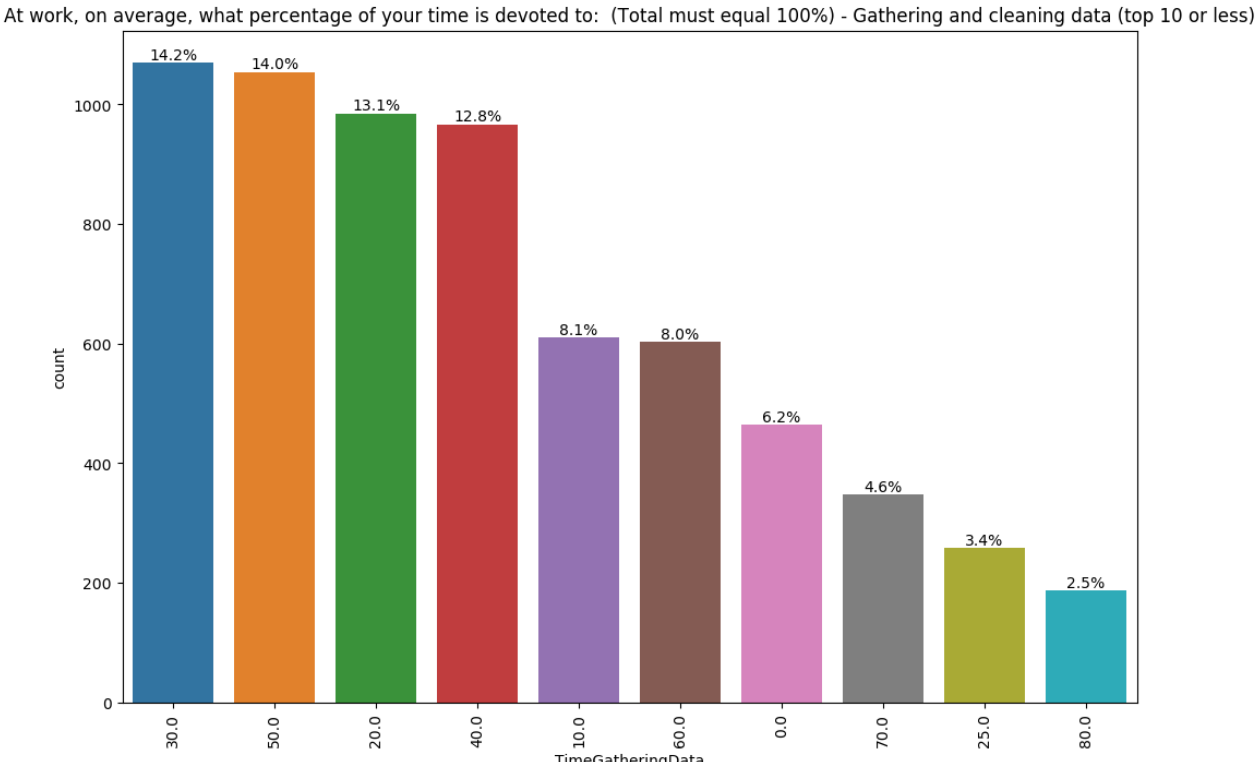
For work, which data science/analytics tools, technologies, and languages have you used in the past year? (Select all that apply) - Selected Choice (top 10 or less)

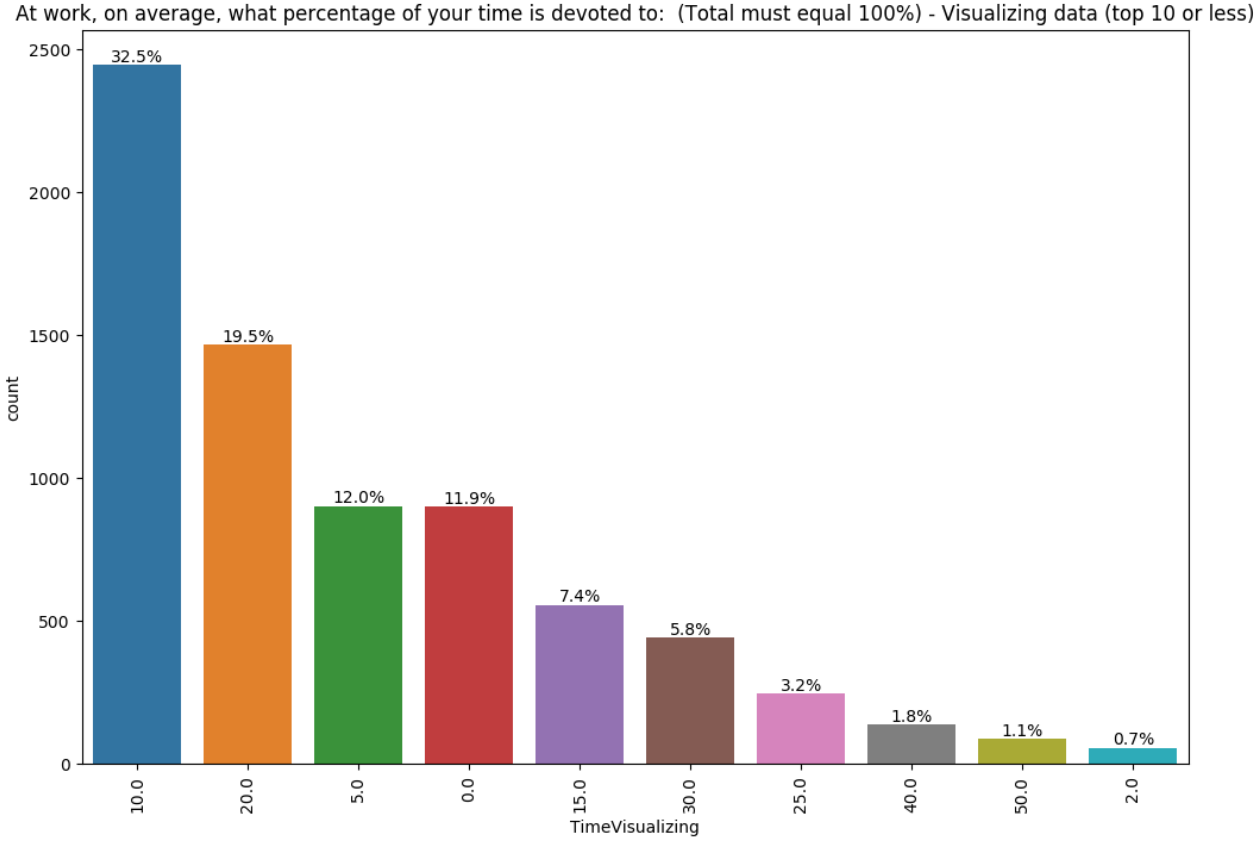


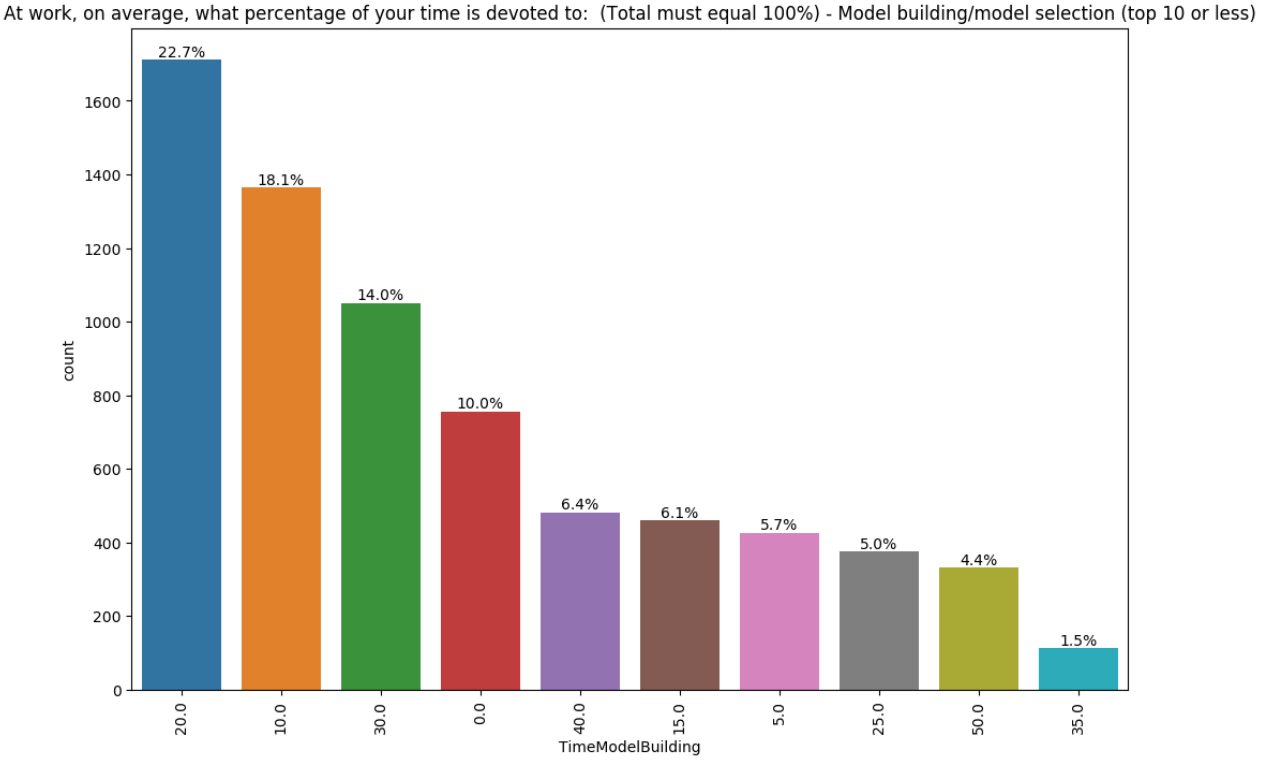
At work, which data science methods do you use? (Select all that apply) - Selected Choice (top 10 or less)

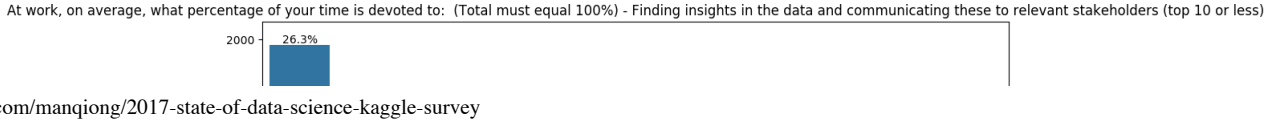


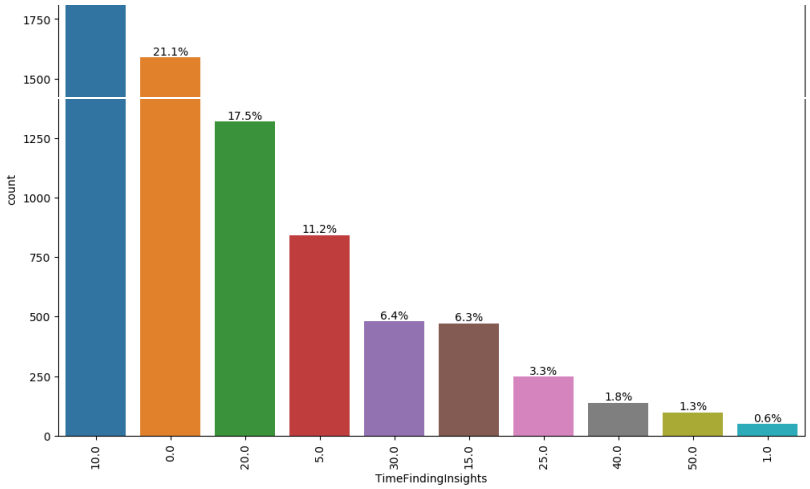

count
0.8%
0.6%
0.6%
0.5%
0.5%
0.5%
0.4%
0.3%
Other
Data Visualization
Logistic Regression
Time Series Analysis
Neural Networks
A/B Testing
Data Visualization,Time Series Analysis
Text Analytics
Decision Trees
CNNs
WorkMethodsSelect

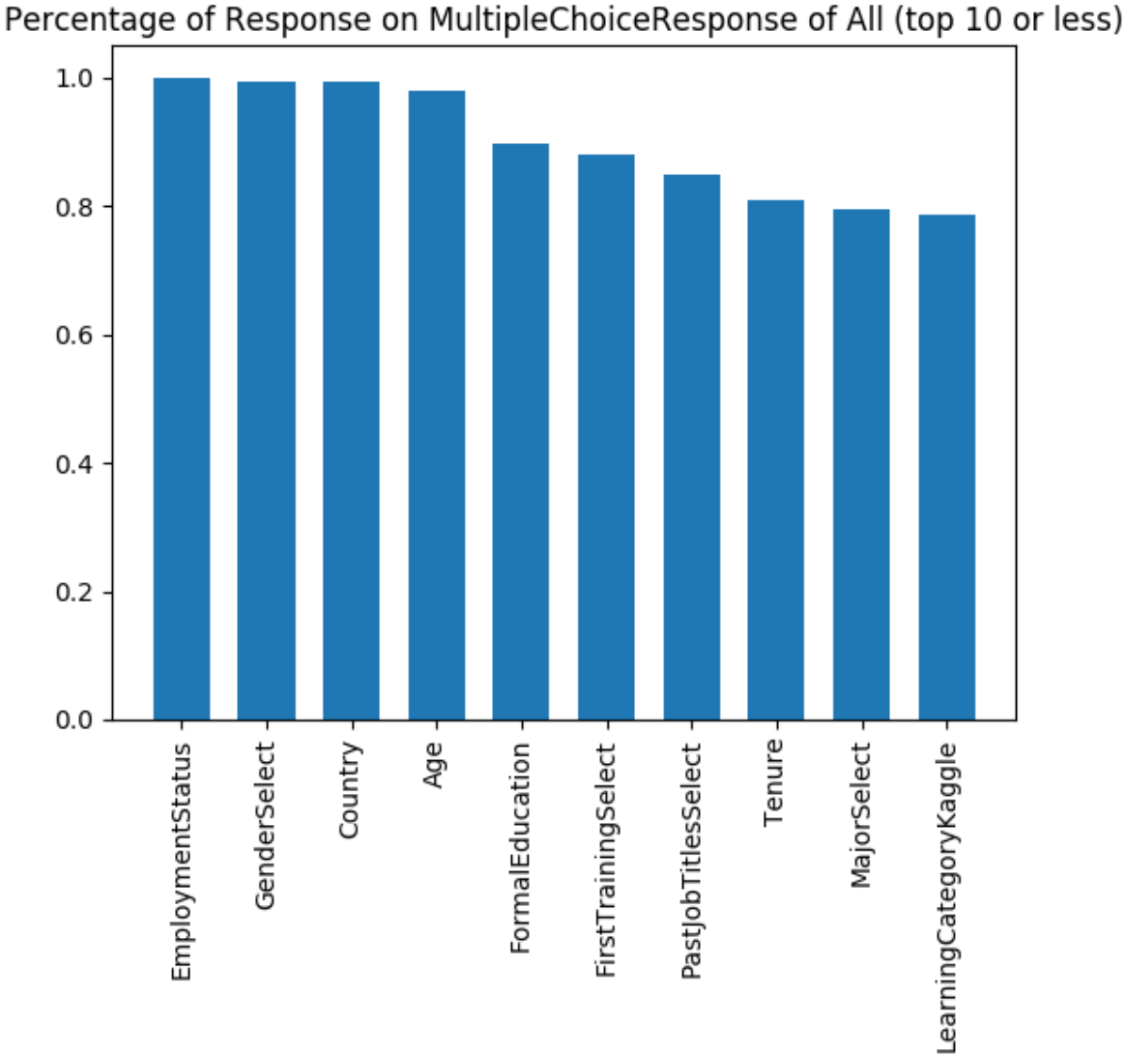

At work, on average, what percentage of your time is devoted to: (Total must equal 100%) - Gathering and cleaning data (top 10 or less)
count
14.2%
14.0%
13.1%
12.8%
8.1%
8.0%
6.2%
4.6%
3.4%
2.5%
30.0
50.0
20.0
40.0
10.0
60.0
0.0
70.0
25.0
80.0
TimeGatheringData

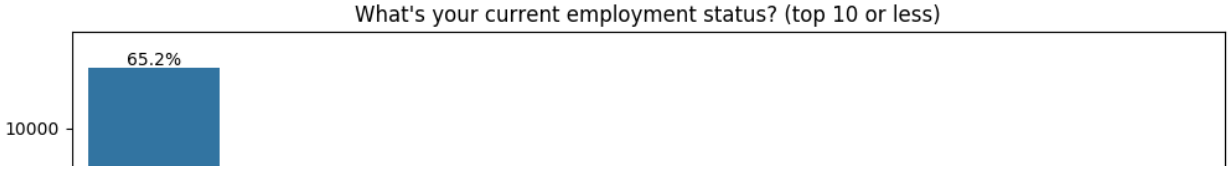

At work, on average, what percentage of your time is devoted to: (Total must equal 100%) - Visualizing data (top 10 or less)
count
TimeVisualizing
32.5%
19.5%
12.0%
11.9%
7.4%
5.8%
3.2%
1.8%
1.1%
0.7%
10.0
20.0
5.0
0.0
15.0
30.0
25.0
40.0
50.0
2.0

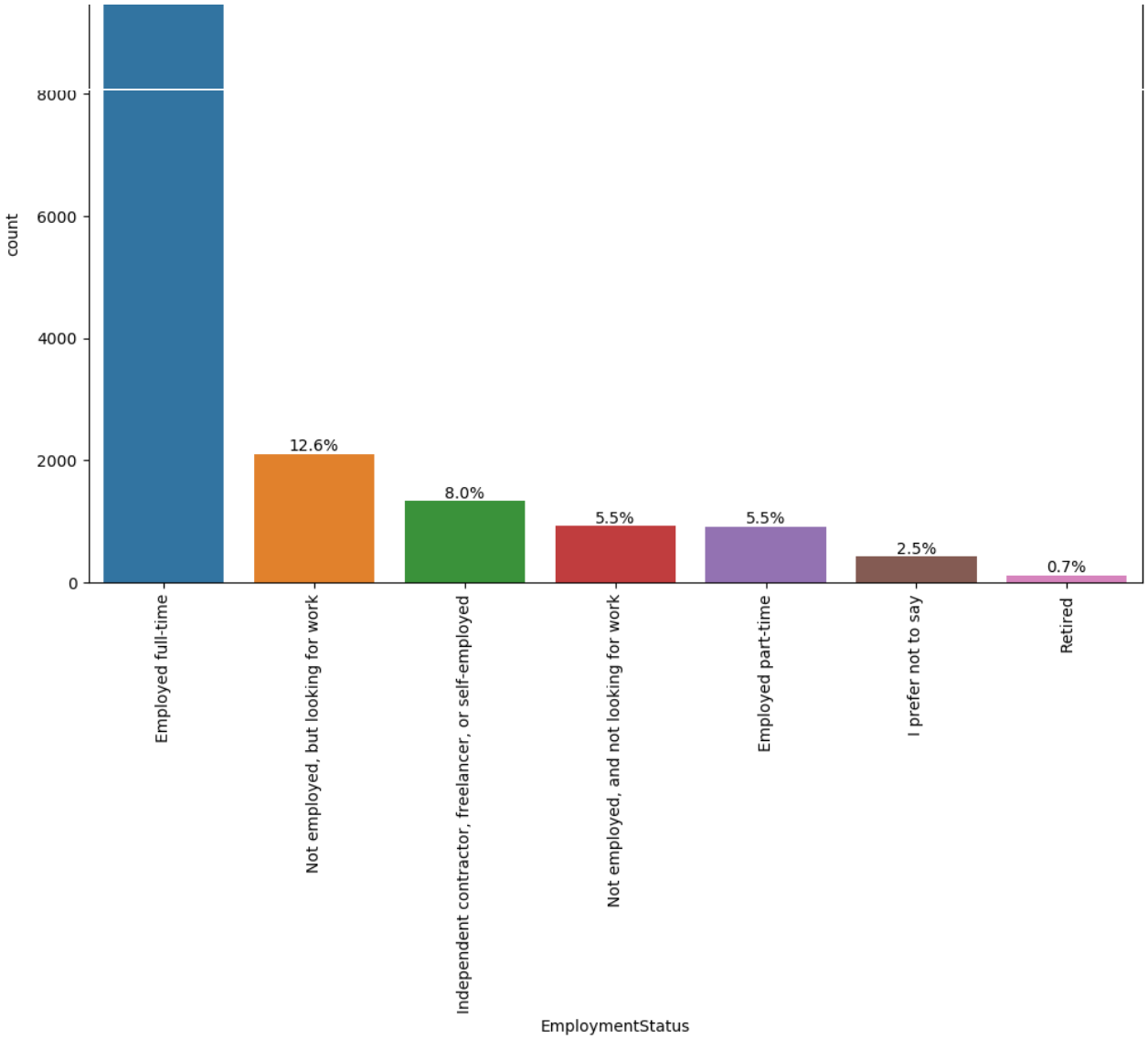

At work, on average, what percentage of your time is devoted to: (Total must equal 100%) - Model building/model selection (top 10 or less)
count
TimeModelBuilding
22.7%
18.1%
14.0%
10.0%
6.4%
6.1%
5.7%
5.0%
4.4%
1.5%
20.0
10.0
30.0
0.0
40.0
15.0
5.0
25.0
50.0
35.0

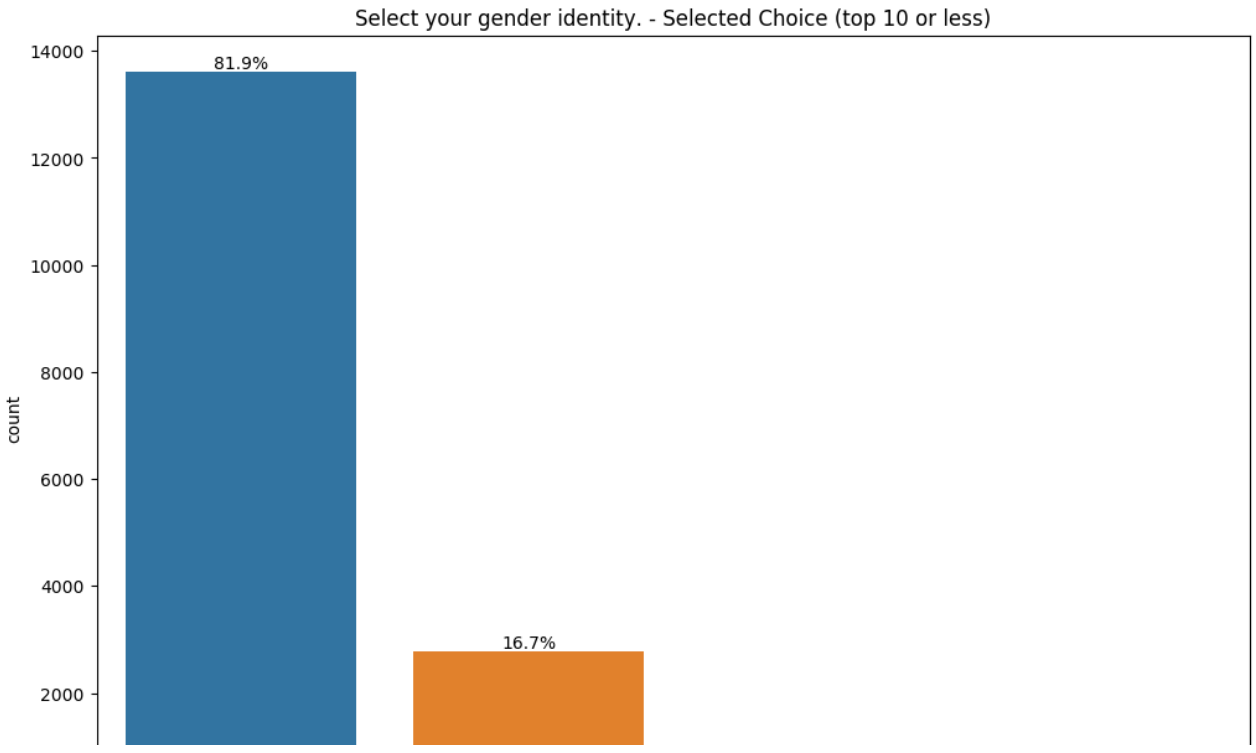

At work, on average, what percentage of your time is devoted to: (Total must equal 100%) - Finding insights in the data and communicating these to relevant stakeholders (top 10 or less)
26.3%

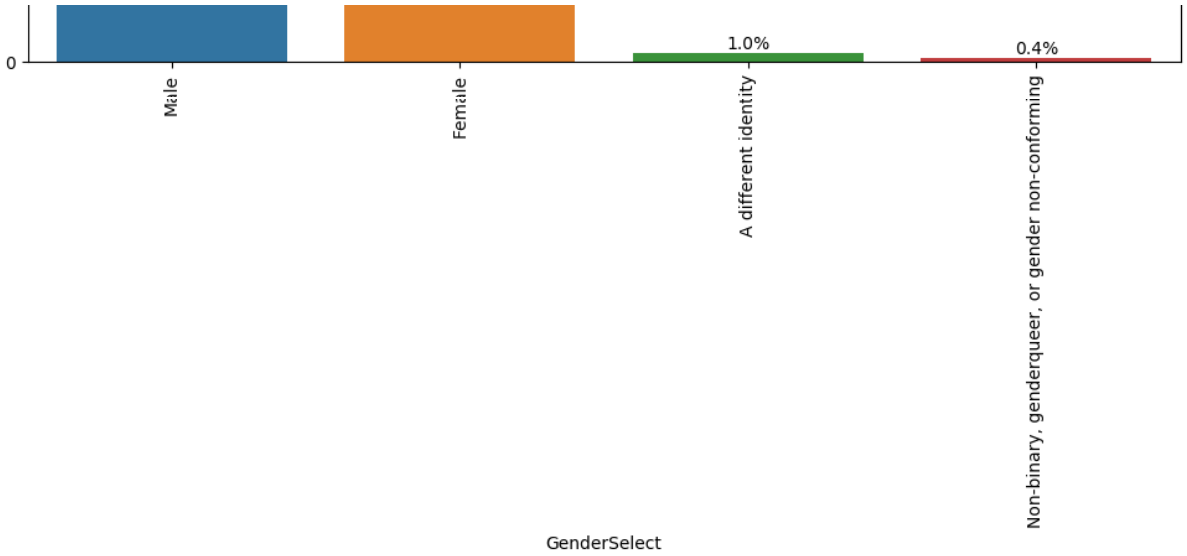

count
TimeFindingInsights
10.0
0.0
20.0
5.0
30.0
15.0
25.0
40.0
50.0
1.0
21.1%
17.5%
11.2%
6.4%
6.3%
3.3%
1.8%
1.3%
0.6%

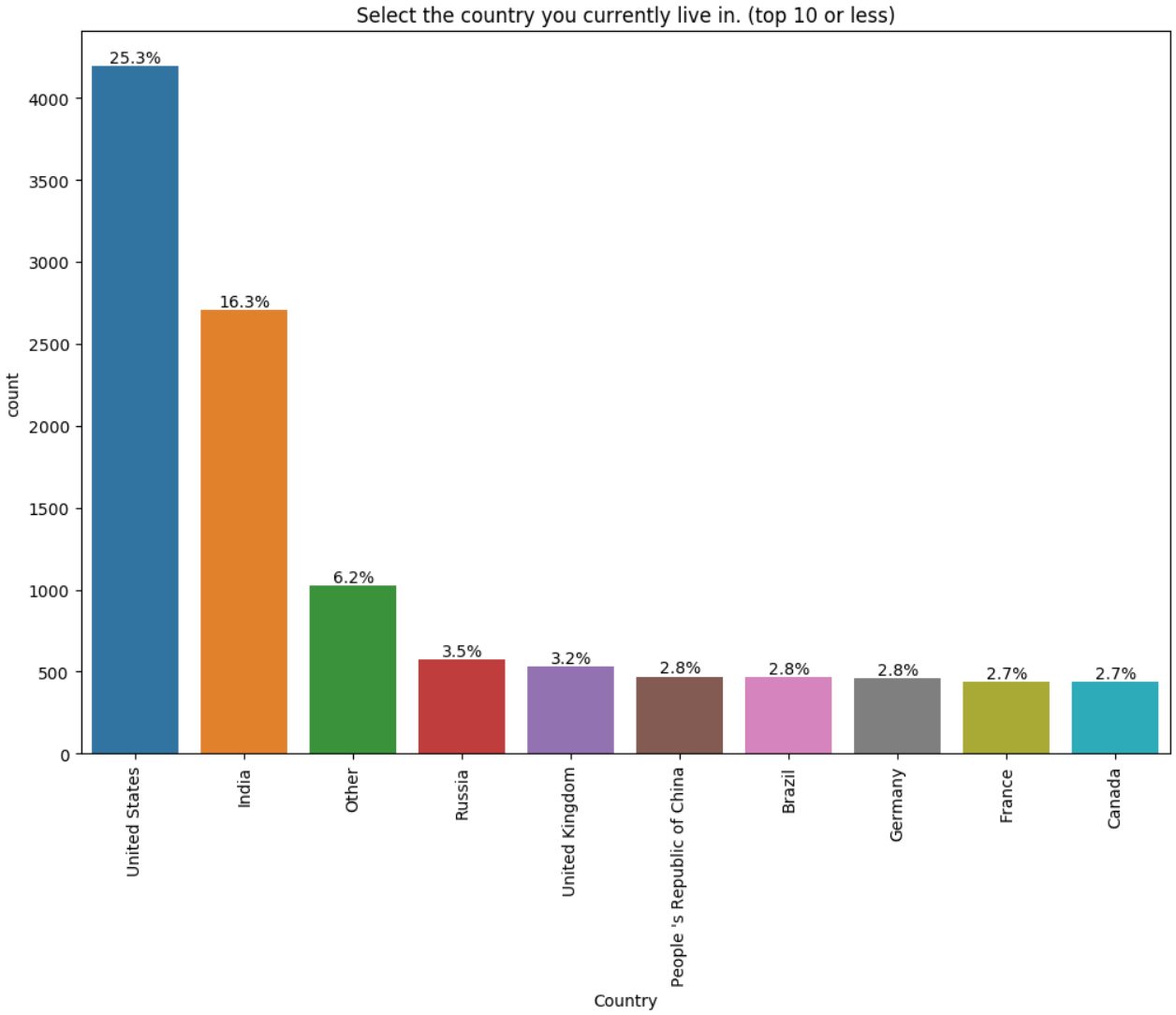

Percentage of Response on MultipleChoiceResponse of All (top 10 or less)
EmploymentStatus
GenderSelect
Country
Age
FormalEducation
FirstTrainingSelect
PastJobTitlesSelect
Tenure
MajorSelect
LearningCategoryKaggle

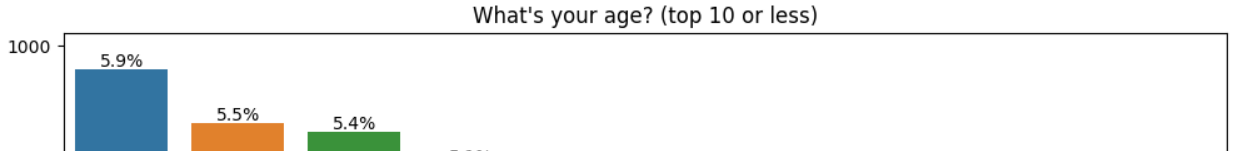

What's your current employment status? (top 10 or less)
65.2%
10000

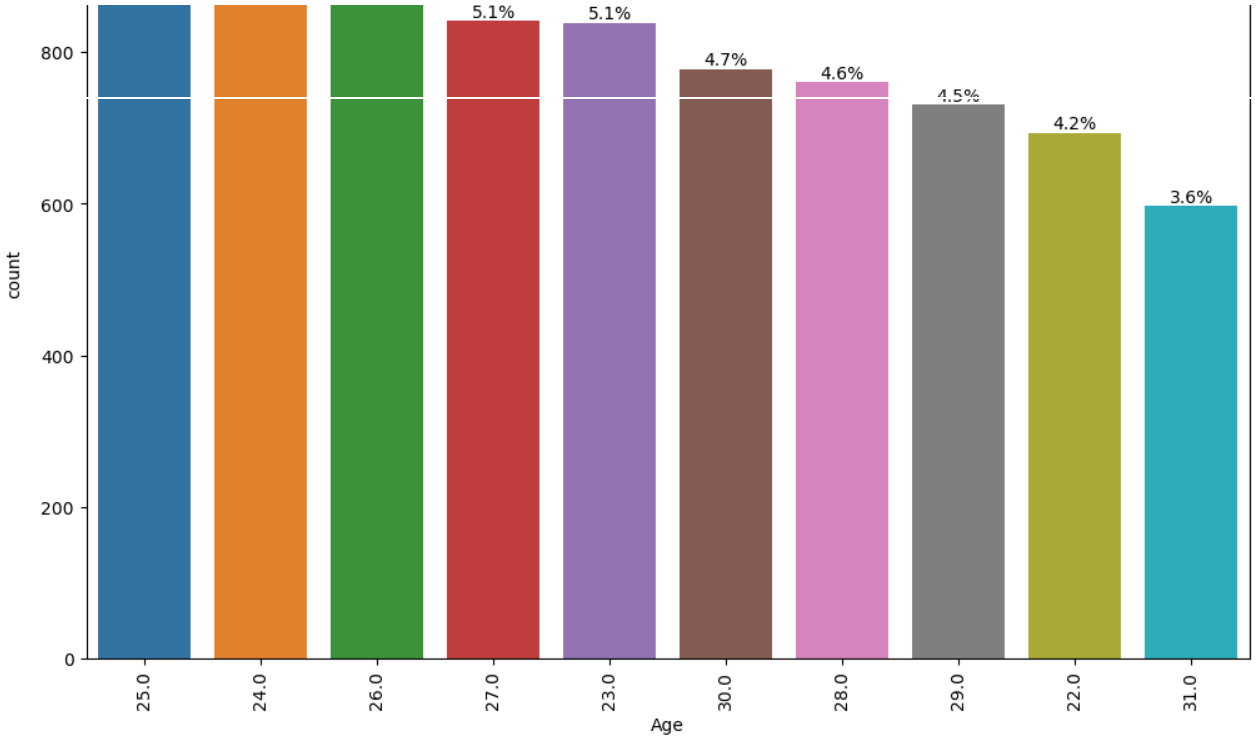

count
8000
6000
4000
2000
0
12.6%
8.0%
5.5%
5.5%
2.5%
0.7%
Employed full-time
Not employed, but looking for work
Independent contractor, freelancer, or self-employed
Not employed, and not looking for work
Employed part-time
I prefer not to say
Retired
EmploymentStatus

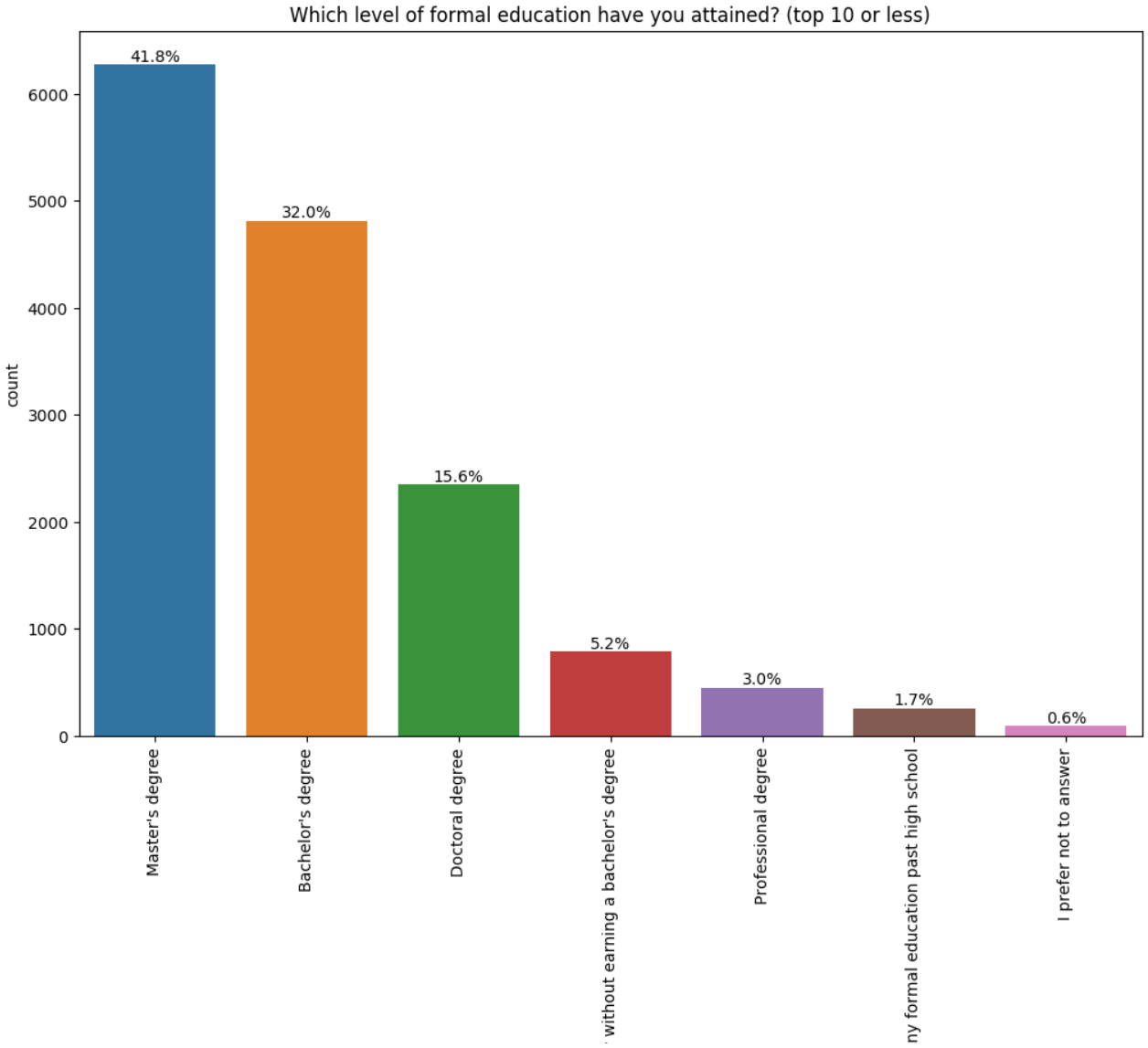

Select your gender identity. - Selected Choice (top 10 or less)
count
14000
12000
10000
8000
6000
4000
2000
81.9%
16.7%


0
1.0%
0.4%
Male
Female
A different identity
Non-binary, genderqueer, or gender non-conforming
GenderSelect

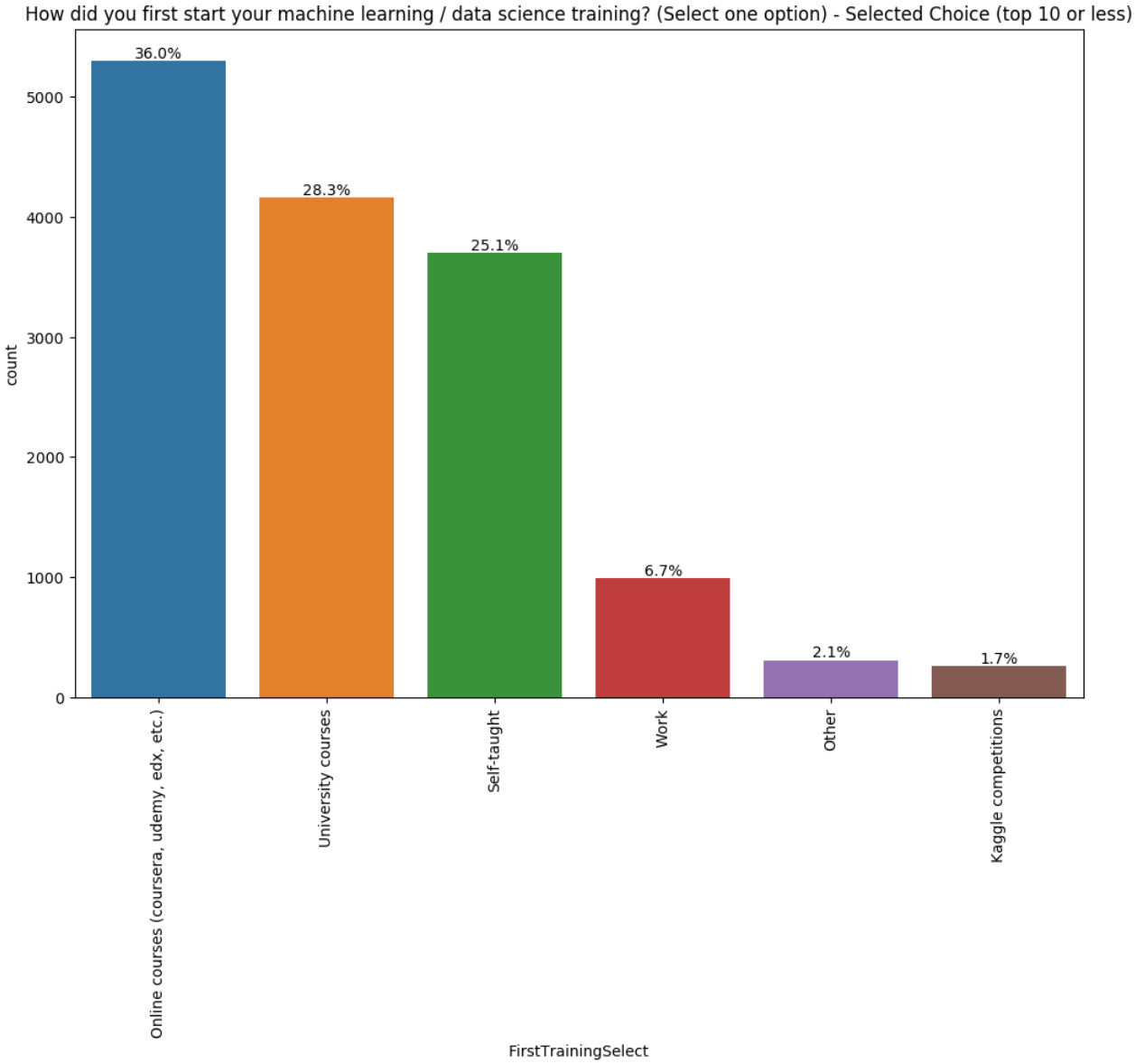

Select the country you currently live in. (top 10 or less)
count
0
500
1000
1500
2000
2500
3000
3500
4000
25.3%
16.3%
6.2%
3.5%
3.2%
2.8%
2.8%
2.8%
2.7%
2.7%
United States
India
Other
Russia
United Kingdom
People 's Republic of China
Brazil
Germany
France
Canada
Country

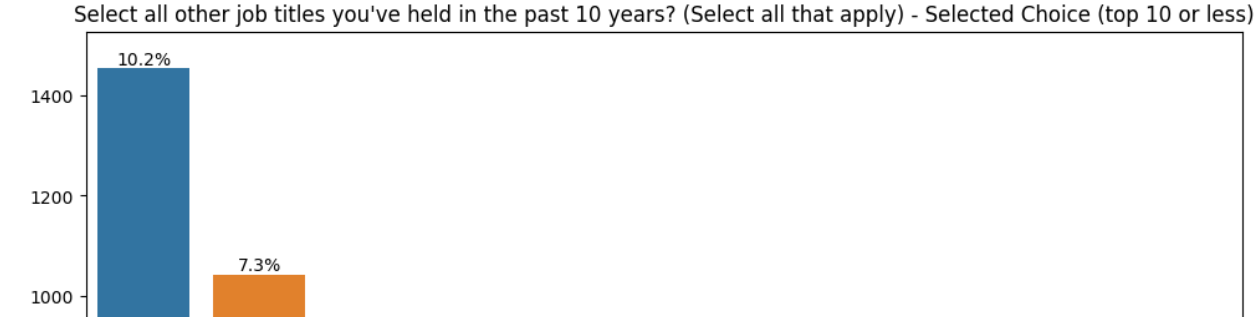

What's your age? (top 10 or less)
1000
5.9%
5.5%
5.4%

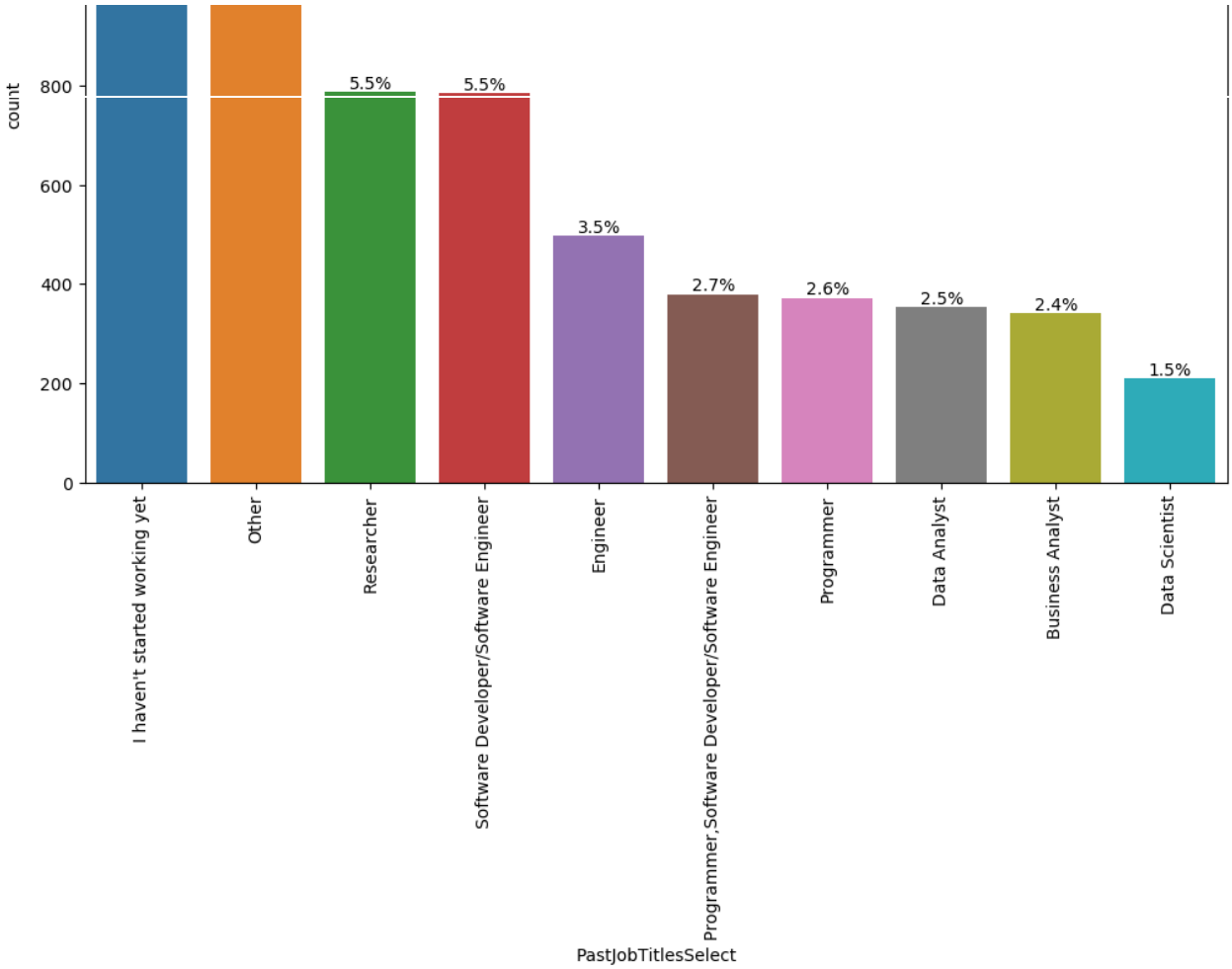

count
800
600
400
200
0
5.1%
5.1%
4.7%
4.6%
4.5%
4.2%
3.6%
25.0
24.0
26.0
27.0
23.0
30.0
28.0
29.0
22.0
31.0
Age

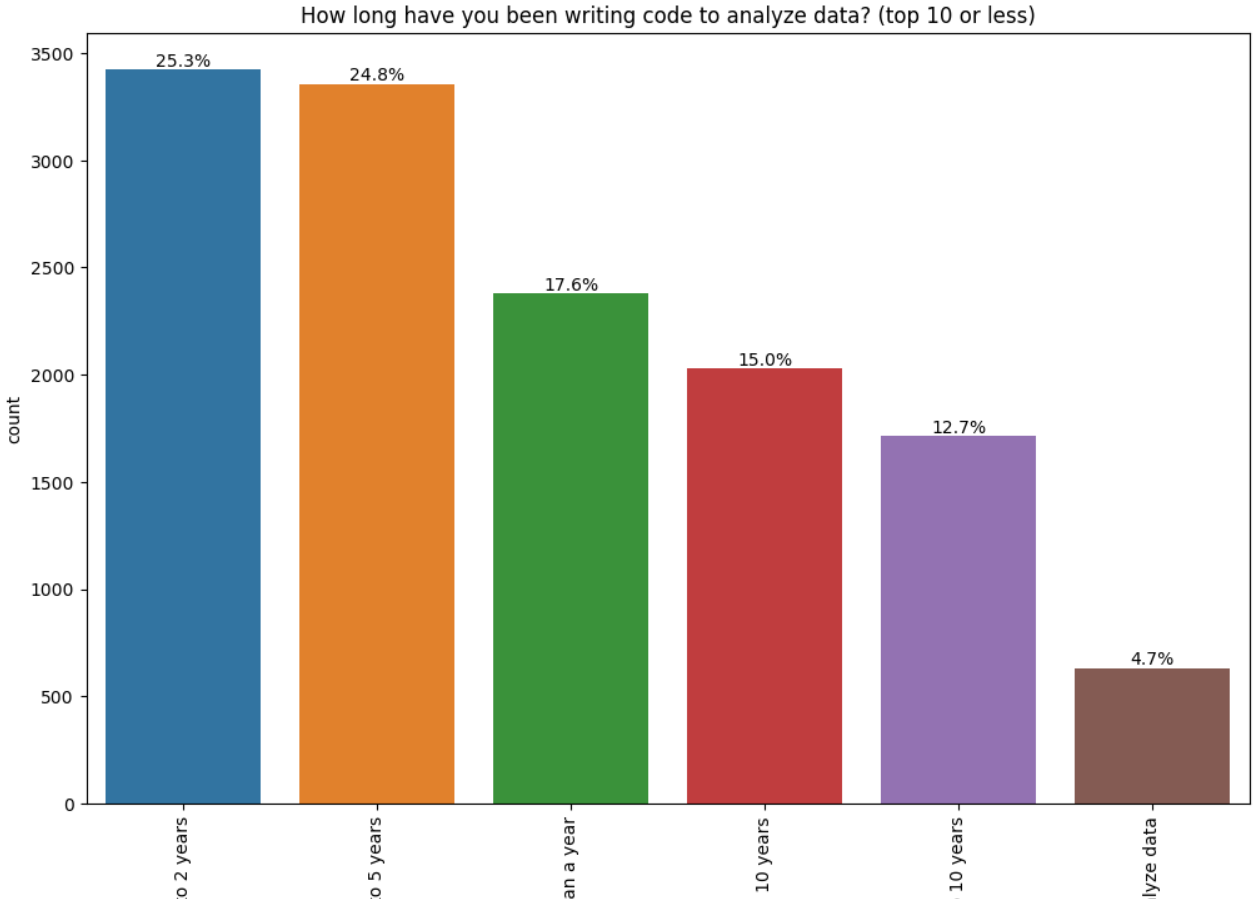

Which level of formal education have you attained? (top 10 or less)
count
6000
5000
4000
3000
2000
1000
0
41.8%
32.0%
15.6%
5.2%
3.0%
1.7%
0.6%
Master's degree
Bachelor's degree
Doctoral degree
without earning a bachelor's degree
Professional degree
ny formal education past high school
I prefer not to answer

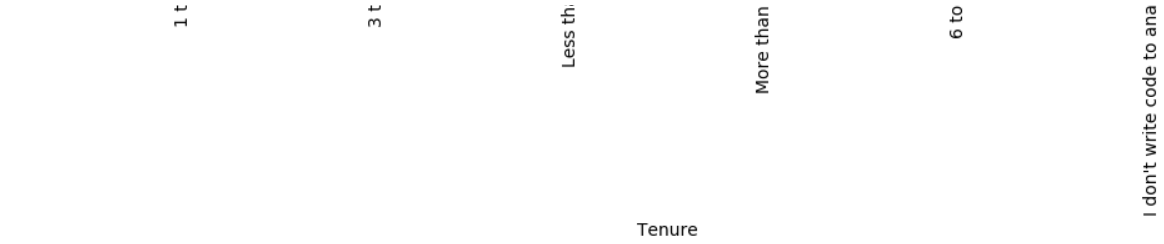

Some college/university study
I did not complete a
FormalEducation

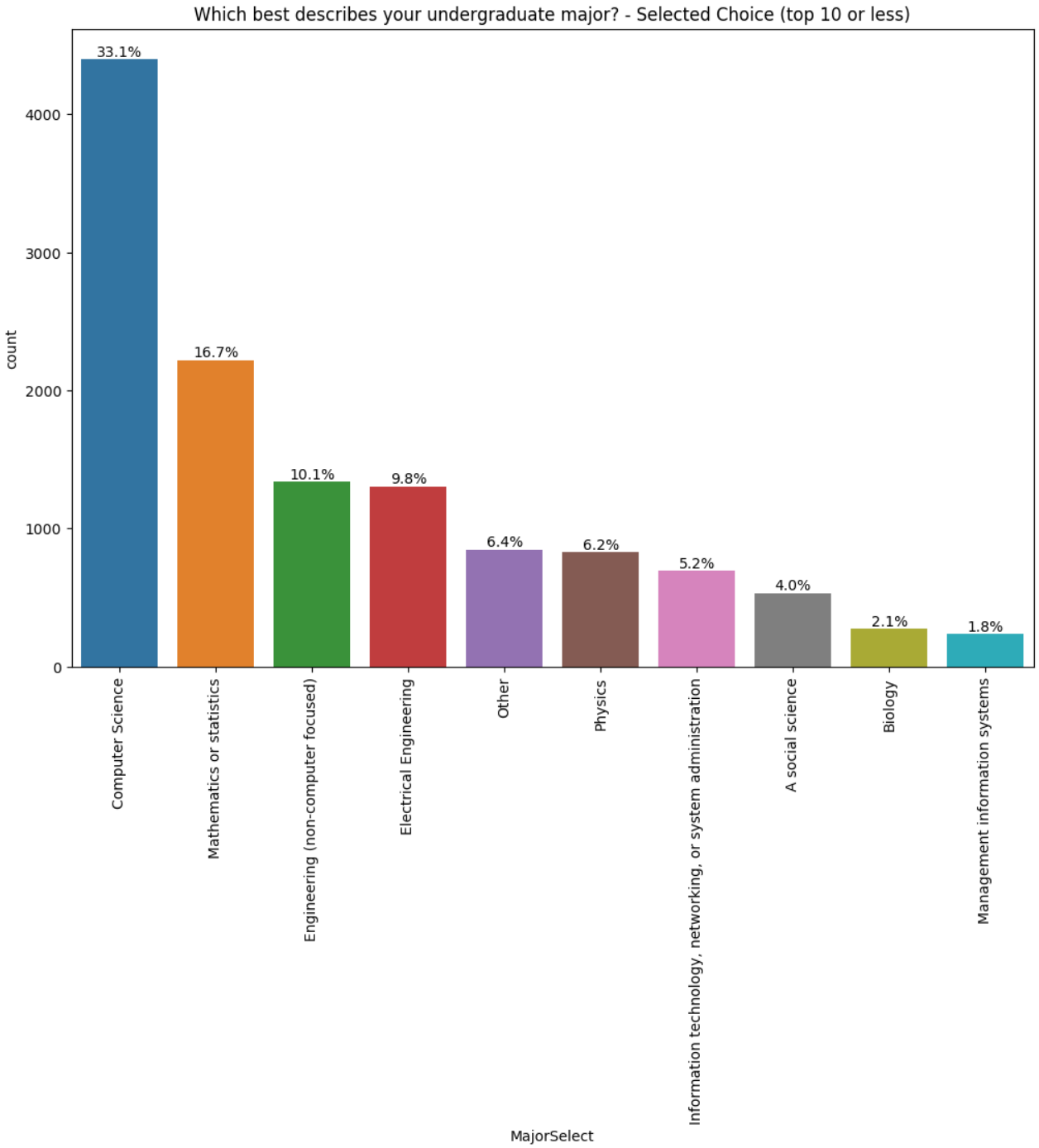

How did you first start your machine learning / data science training? (Select one option) - Selected Choice (top 10 or less)
count
0
1000
2000
3000
4000
5000
36.0%
28.3%
25.1%
6.7%
2.1%
1.7%
Online courses (coursera, udemy, edx, etc.)
University courses
Self-taught
Work
Other
Kaggle competitions
FirstTrainingSelect

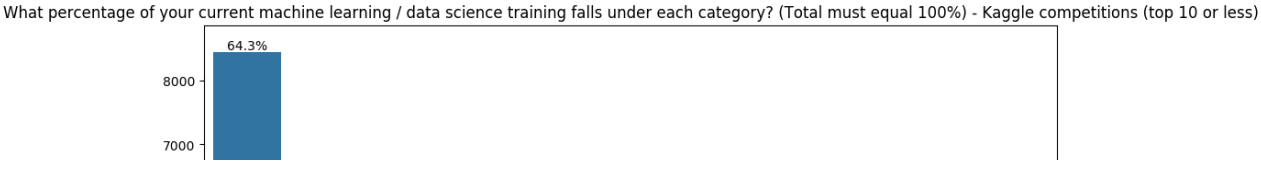

Select all other job titles you've held in the past 10 years? (Select all that apply) - Selected Choice (top 10 or less)
1000
1200
1400
10.2%
7.3%

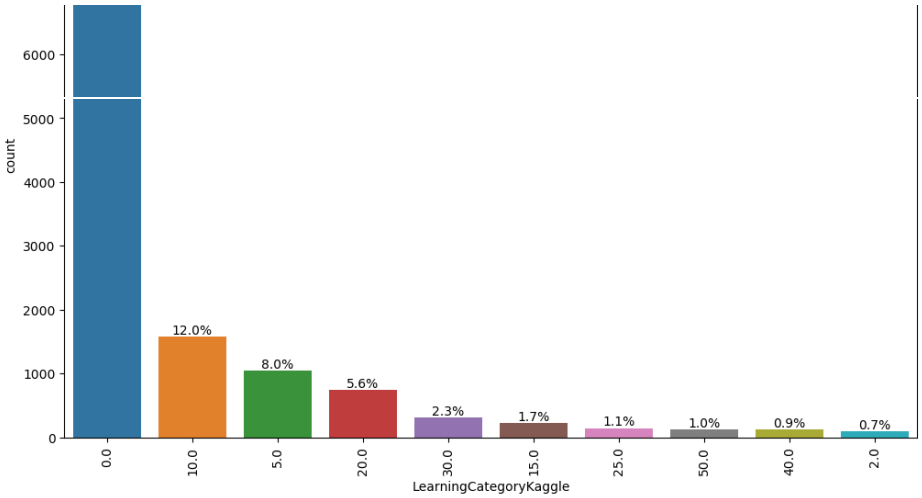

count
800
600
400
200
0
5.5%
5.5%
3.5%
2.7%
2.6%
2.5%
2.4%
1.5%
I haven't started working yet
Other
Researcher
Software Developer/Software Engineer
Engineer
Programmer,Software Developer/Software Engineer
Programmer
Data Analyst
Business Analyst
Data Scientist
PastJobTitlesSelect

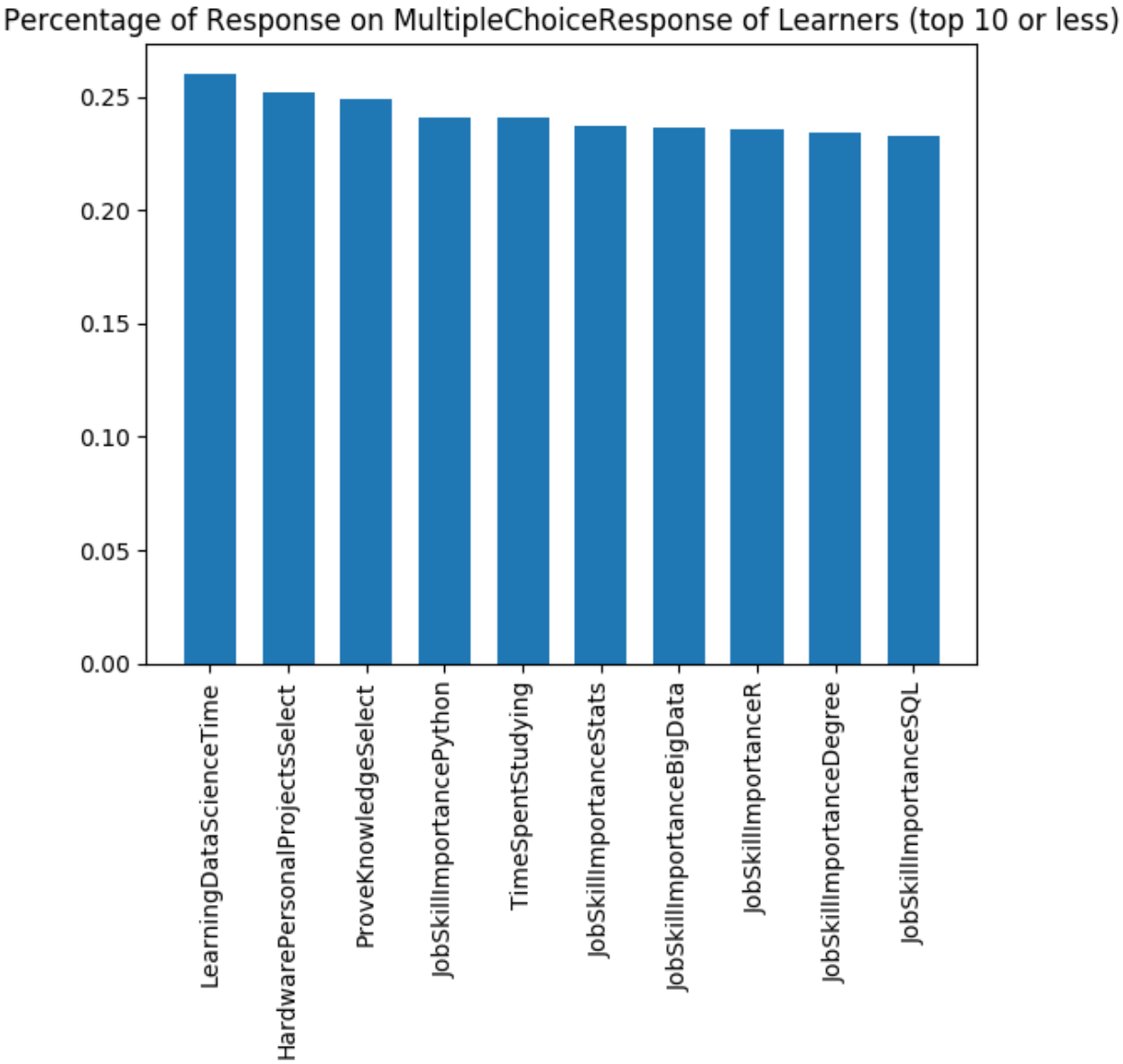

How long have you been writing code to analyze data? (top 10 or less)
count
3500
3000
2500
2000
1500
1000
500
0
25.3%
24.8%
17.6%
15.0%
12.7%
4.7%
o 2 years
o 5 years
an a year
10 years
10 years
lyze data

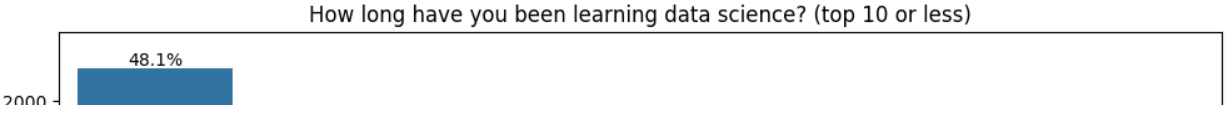

1 t
3 t
Less th
More than
6 to
I don't write code to ana
Tenure

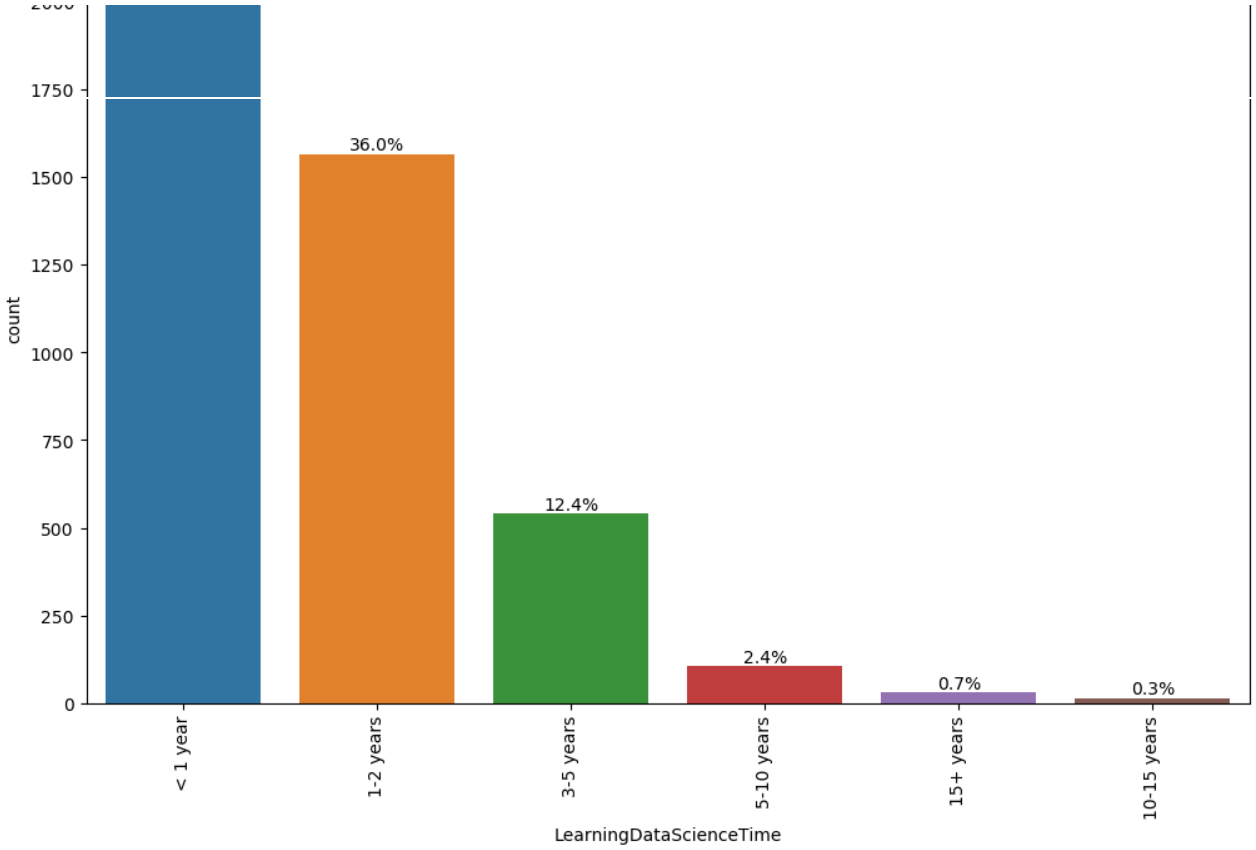

Which best describes your undergraduate major? - Selected Choice (top 10 or less)
count
0
1000
2000
3000
4000
33.1%
16.7%
10.1%
9.8%
6.4%
6.2%
5.2%
4.0%
2.1%
1.8%
Computer Science
Mathematics or statistics
Engineering (non-computer focused)
Electrical Engineering
Other
Physics
Information technology, networking, or system administration
A social science
Biology
Management information systems
MajorSelect

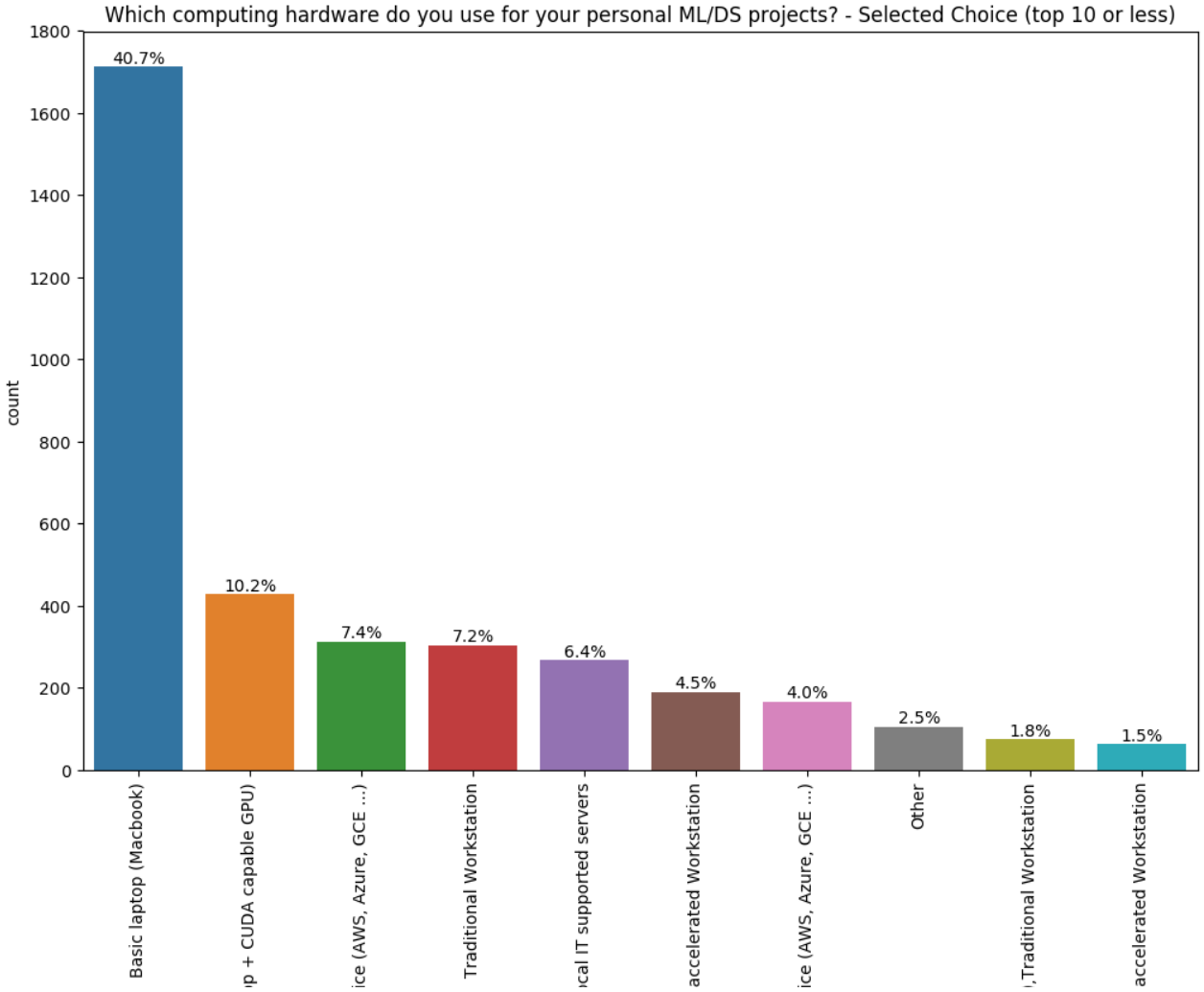

What percentage of your current machine learning / data science training falls under each category? (Total must equal 100%) - Kaggle competitions (top 10 or less)
64.3%
8000
7000

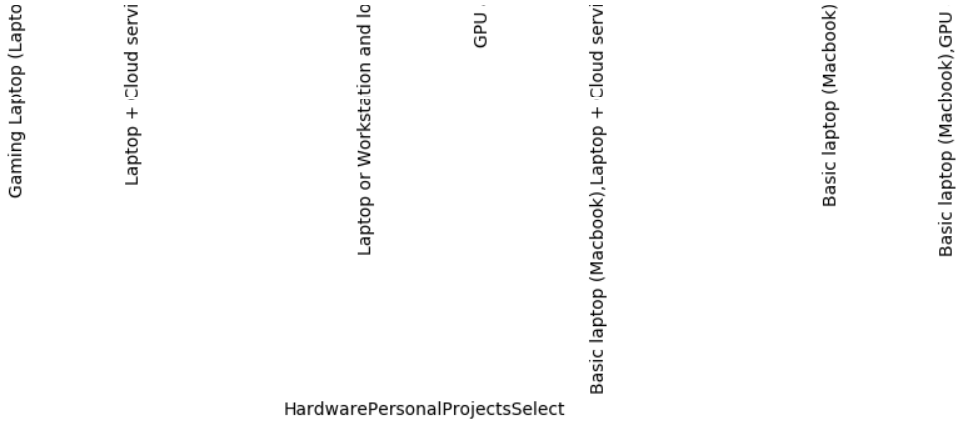

count
6000
5000
4000
3000
2000
1000
0
12.0%
8.0%
5.6%
2.3%
1.7%
1.1%
1.0%
0.9%
0.7%
0.0
10.0
5.0
20.0
30.0
15.0
25.0
50.0
40.0
2.0
LearningCategoryKaggle

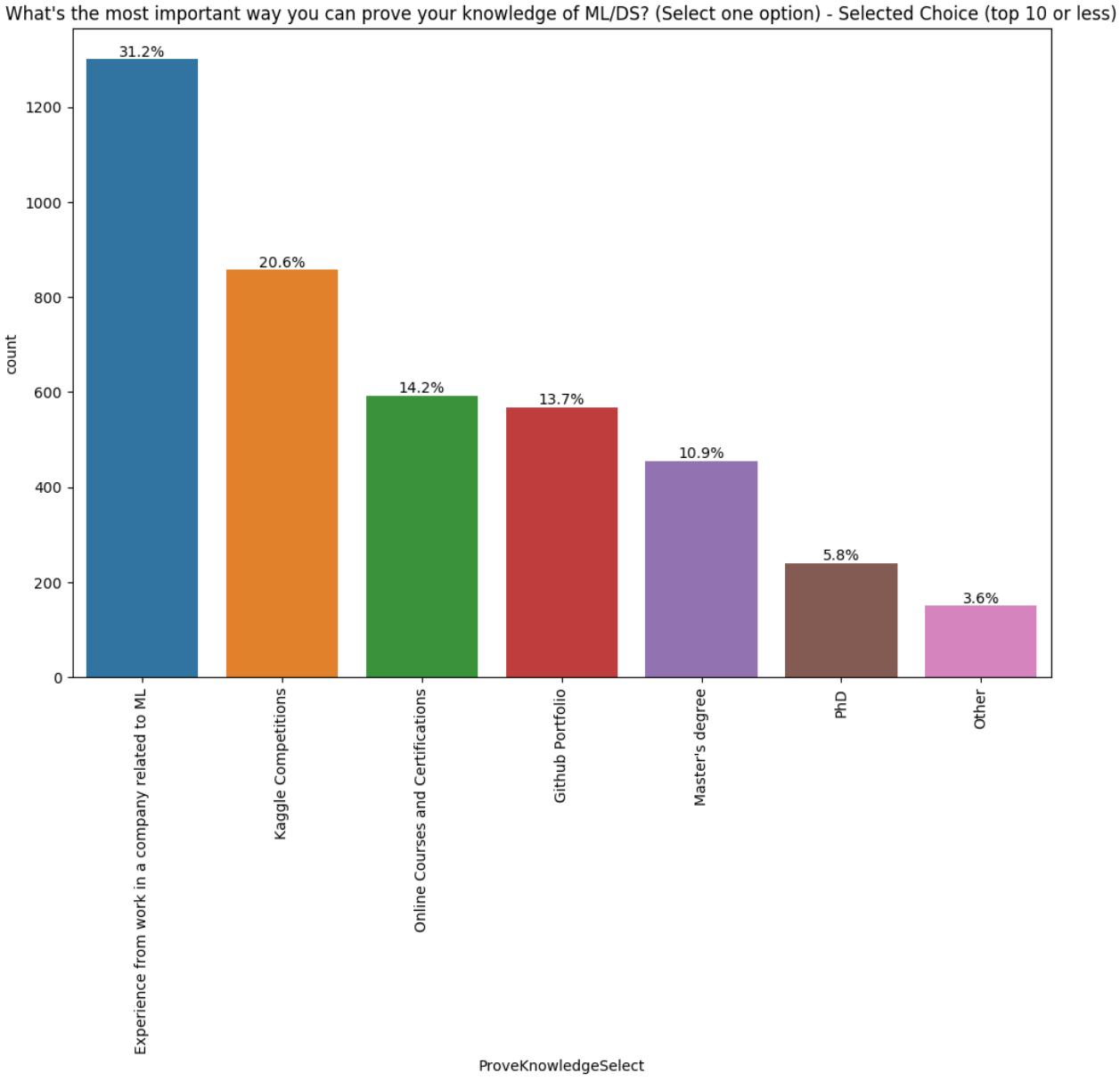

Percentage of Response on MultipleChoiceResponse of Learners (top 10 or less)
0.25
0.20
0.15
0.10
0.05
0.00
LearningDataScienceTime
HardwarePersonalProjectsSelect
ProveKnowledgeSelect
JobSkillImportancePython
TimeSpentStudying
JobSkillImportanceStats
JobSkillImportanceBigData
JobSkillImportanceR
JobSkillImportanceDegree
JobSkillImportanceSQL

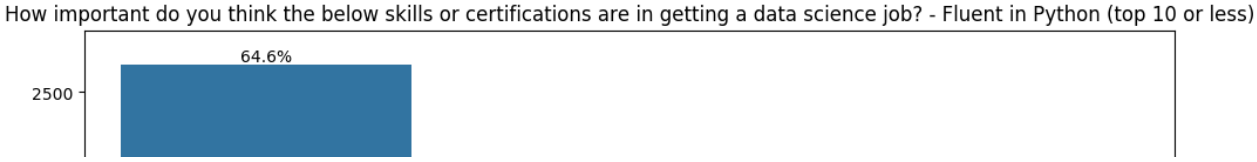

How long have you been learning data science? (top 10 or less)
48.1%
2000

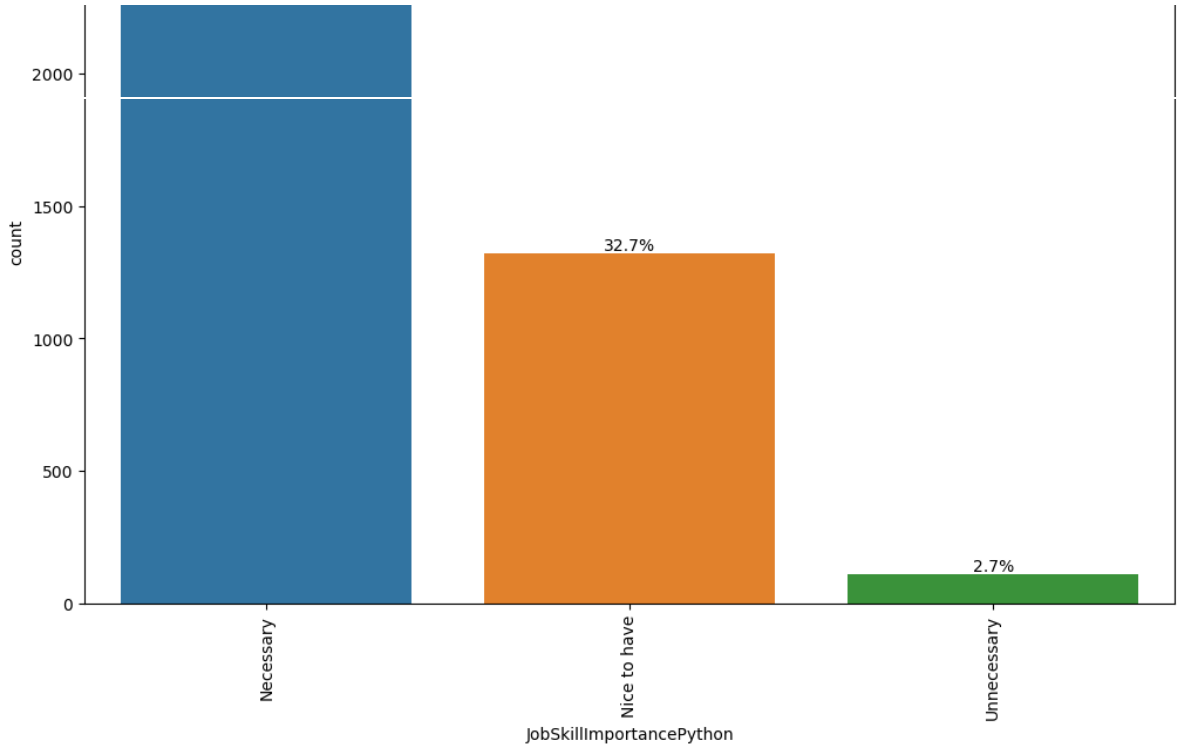

count
1750
1500
1250
1000
750
500
250
0
36.0%
12.4%
2.4%
0.7%
0.3%
< 1 year
1-2 years
3-5 years
5-10 years
15+ years
10-15 years
LearningDataScienceTime

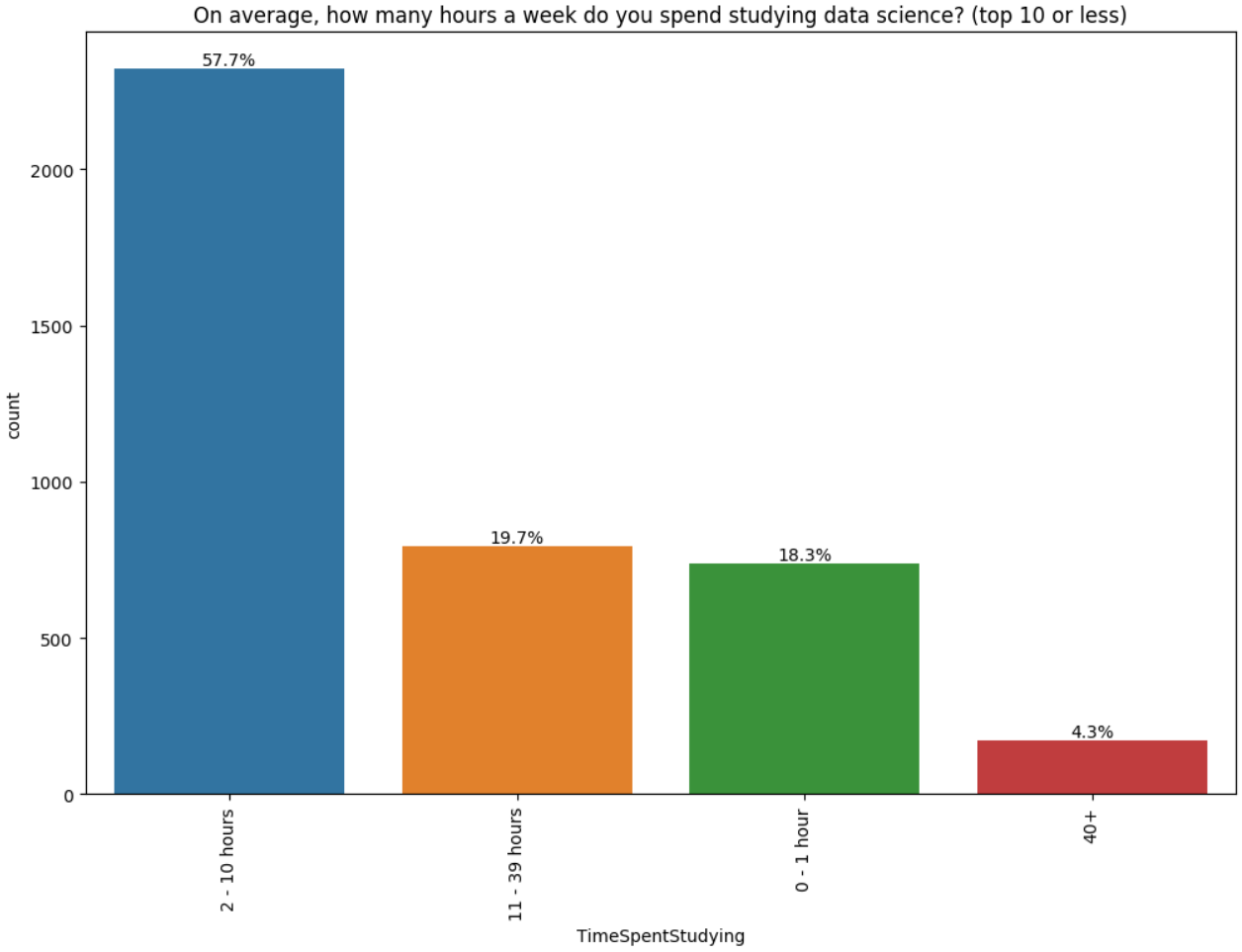

Which computing hardware do you use for your personal ML/DS projects? - Selected Choice (top 10 or less)
count
1800
1600
1400
1200
1000
800
600
400
200
0
40.7%
10.2%
7.4%
7.2%
6.4%
4.5%
4.0%
2.5%
1.8%
1.5%
Basic laptop (Macbook)
Traditional Workstation
Other

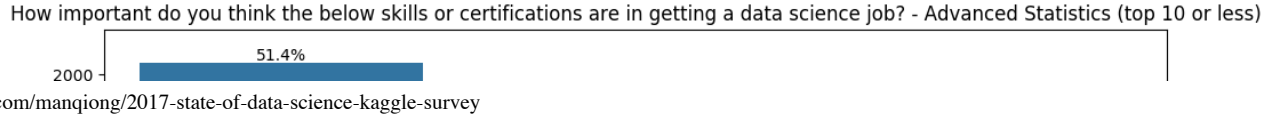

Gaming Laptop (Lapto
Laptop + Cloud servi
Laptop or Workstation and lo
GPU
Basic laptop (Macbook),Laptop + Cloud servi
Basic laptop (Macbook)
Basic laptop (Macbook),GPU
HardwarePersonalProjectsSelect

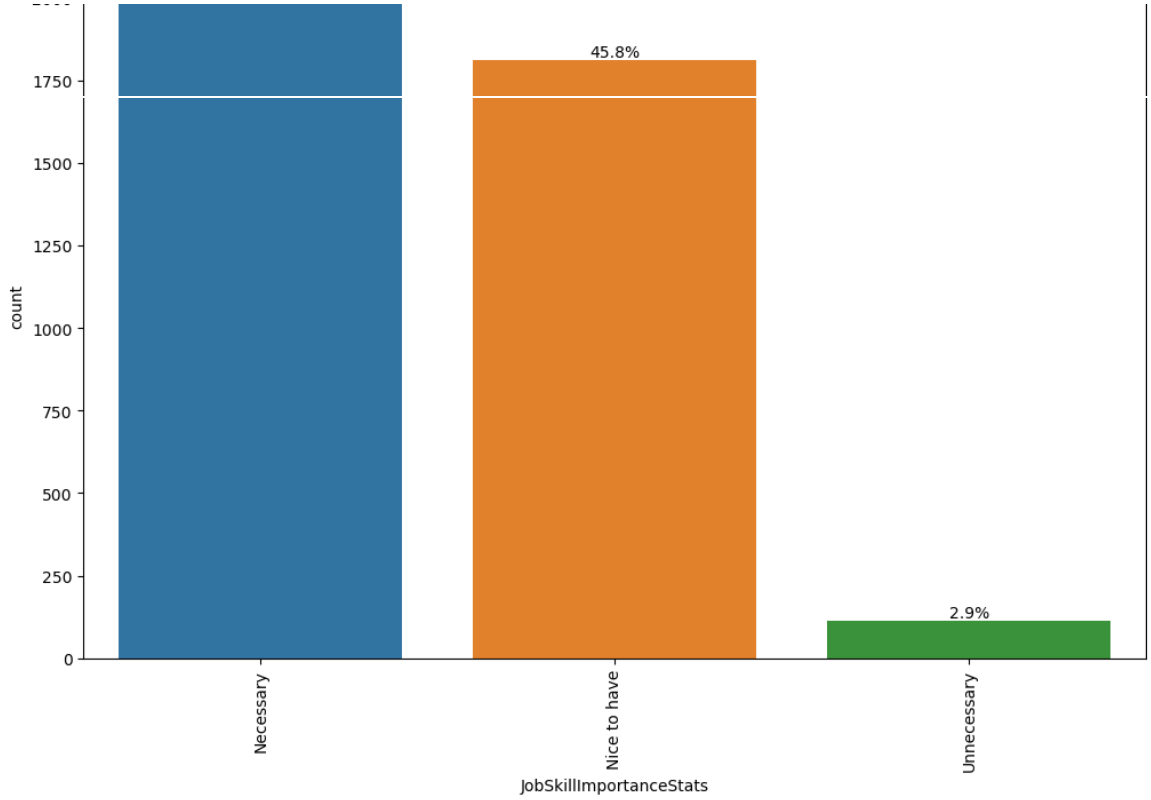

What's the most important way you can prove your knowledge of ML/DS? (Select one option) - Selected Choice (top 10 or less)
count
0
200
400
600
800
1000
1200
31.2%
20.6%
14.2%
13.7%
10.9%
5.8%
3.6%
Experience from work in a company related to ML
Kaggle Competitions
Online Courses and Certifications
Github Portfolio
Master's degree
PhD
Other
ProveKnowledgeSelect


How important do you think the below skills or certifications are in getting a data science job? - Fluent in Python (top 10 or less)
2500
64.6%

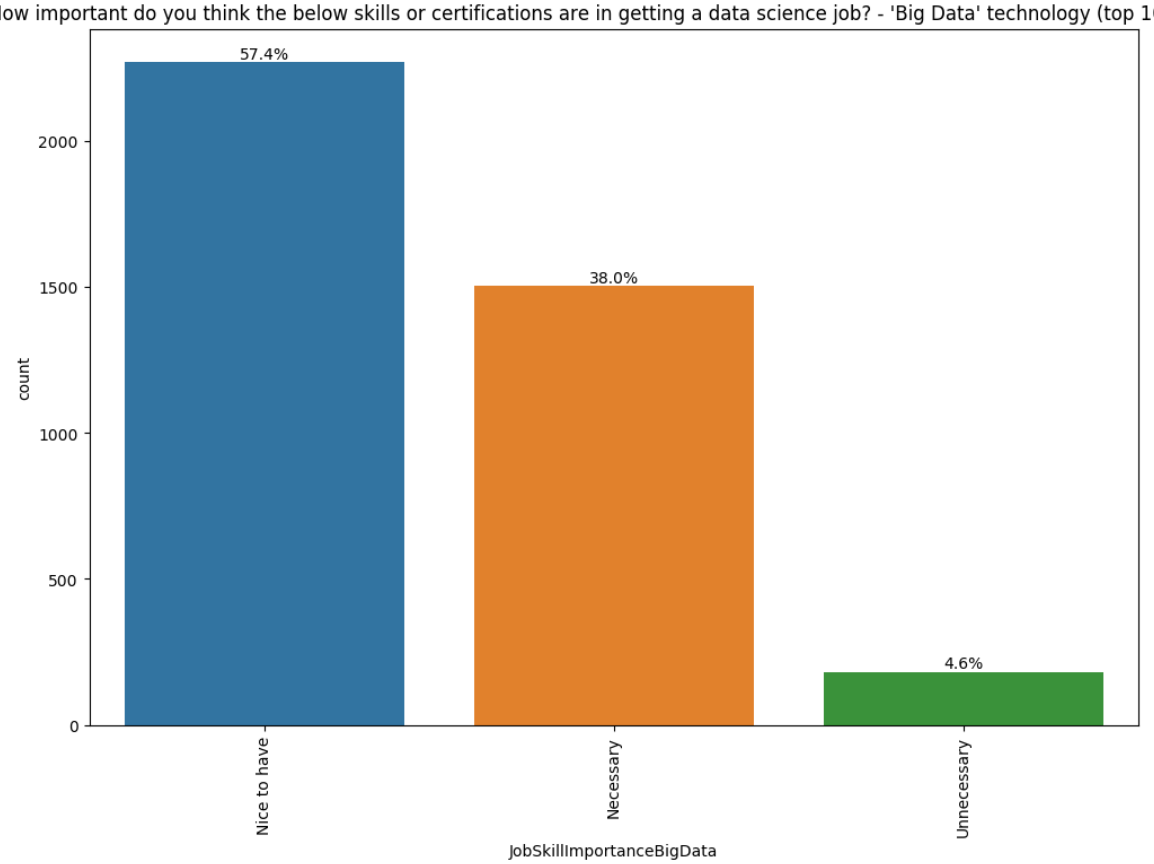

2000
1500
1000
500
0
count
32.7%
2.7%
Necessary
Nice to have
Unnecessary
JobSkillImportancePython


On average, how many hours a week do you spend studying data science? (top 10 or less)
57.7%
19.7%
18.3%
4.3%
2000
1500
1000
500
0
count
2 - 10 hours
11 - 39 hours
0 - 1 hour
40+
TimeSpentStudying

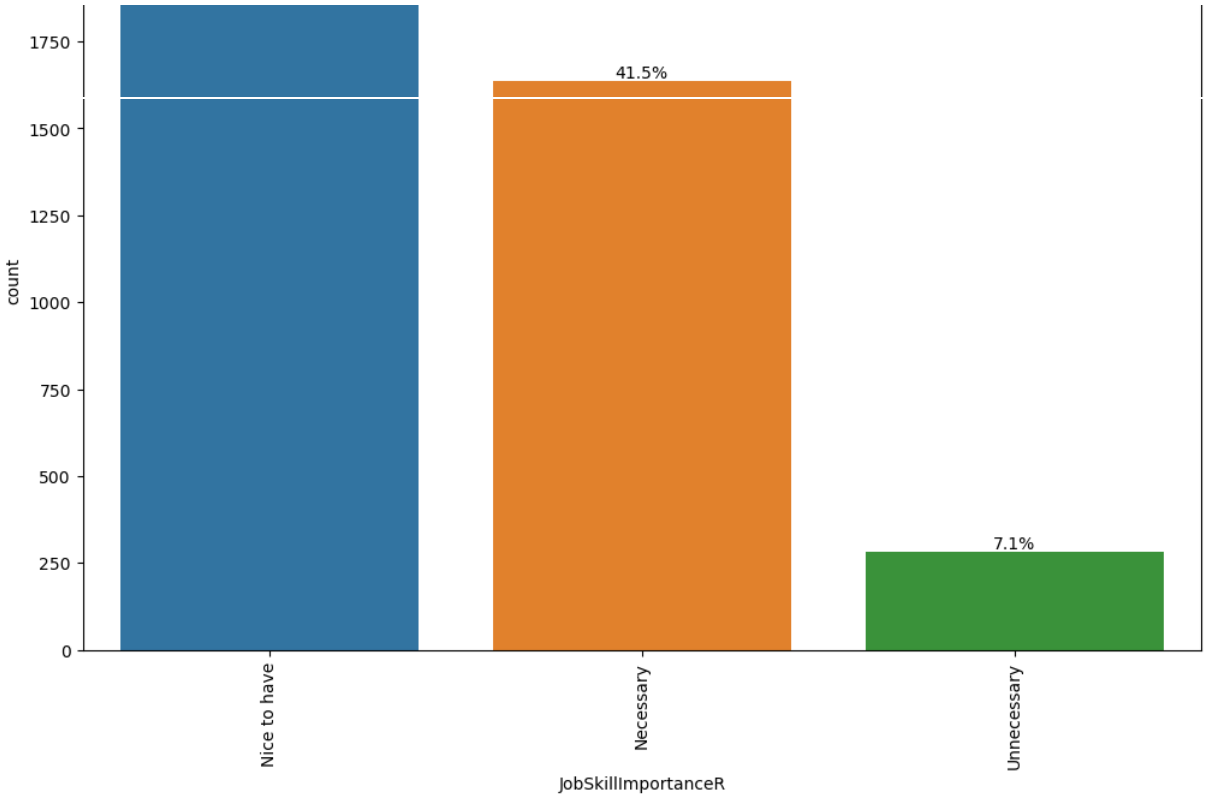

How important do you think the below skills or certifications are in getting a data science job? - Advanced Statistics (top 10 or less)
51.4%
2000



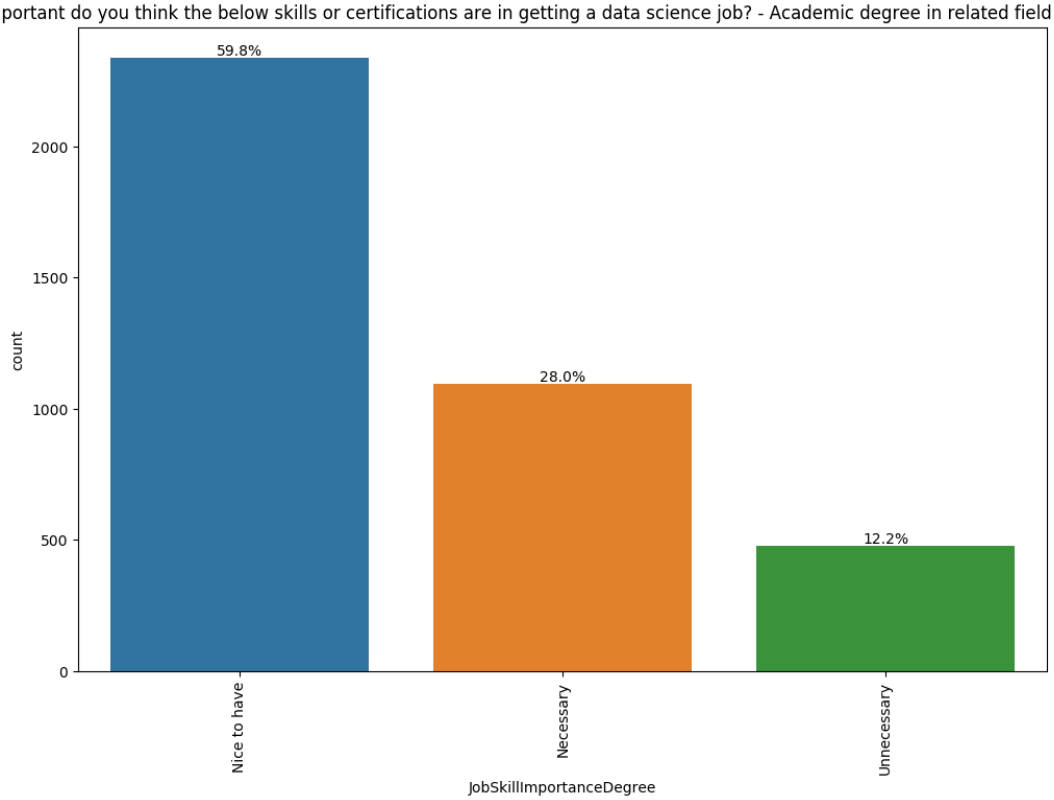
How important do you think the below skills or certifications are in getting a data science job? - 'Big Data' technology (top 10 or less)



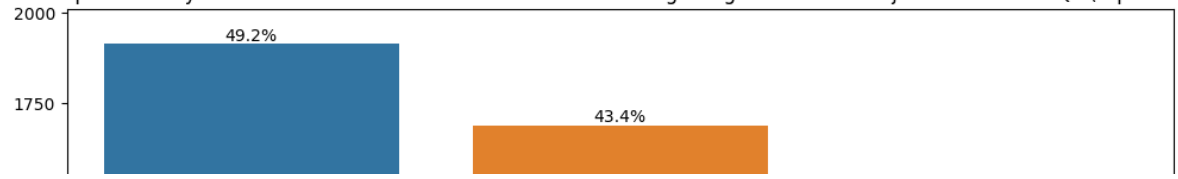
How important do you think the below skills or certifications are in getting a data science job? - Fluent in R (top 10 or less)

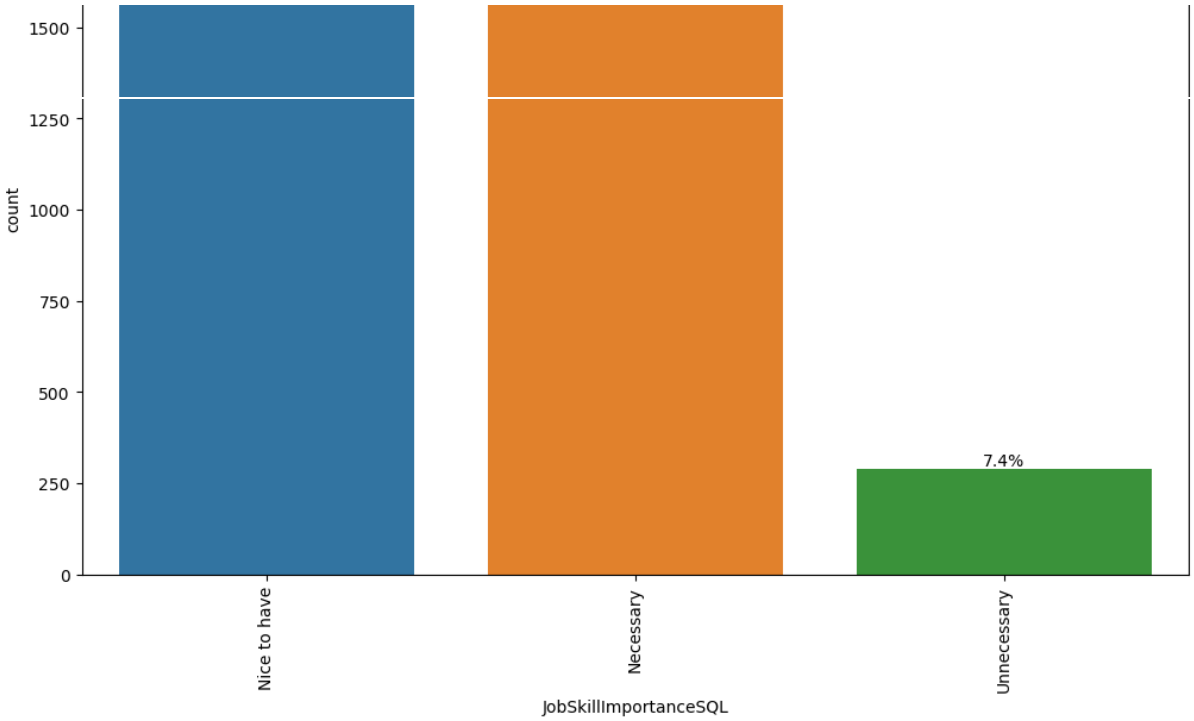


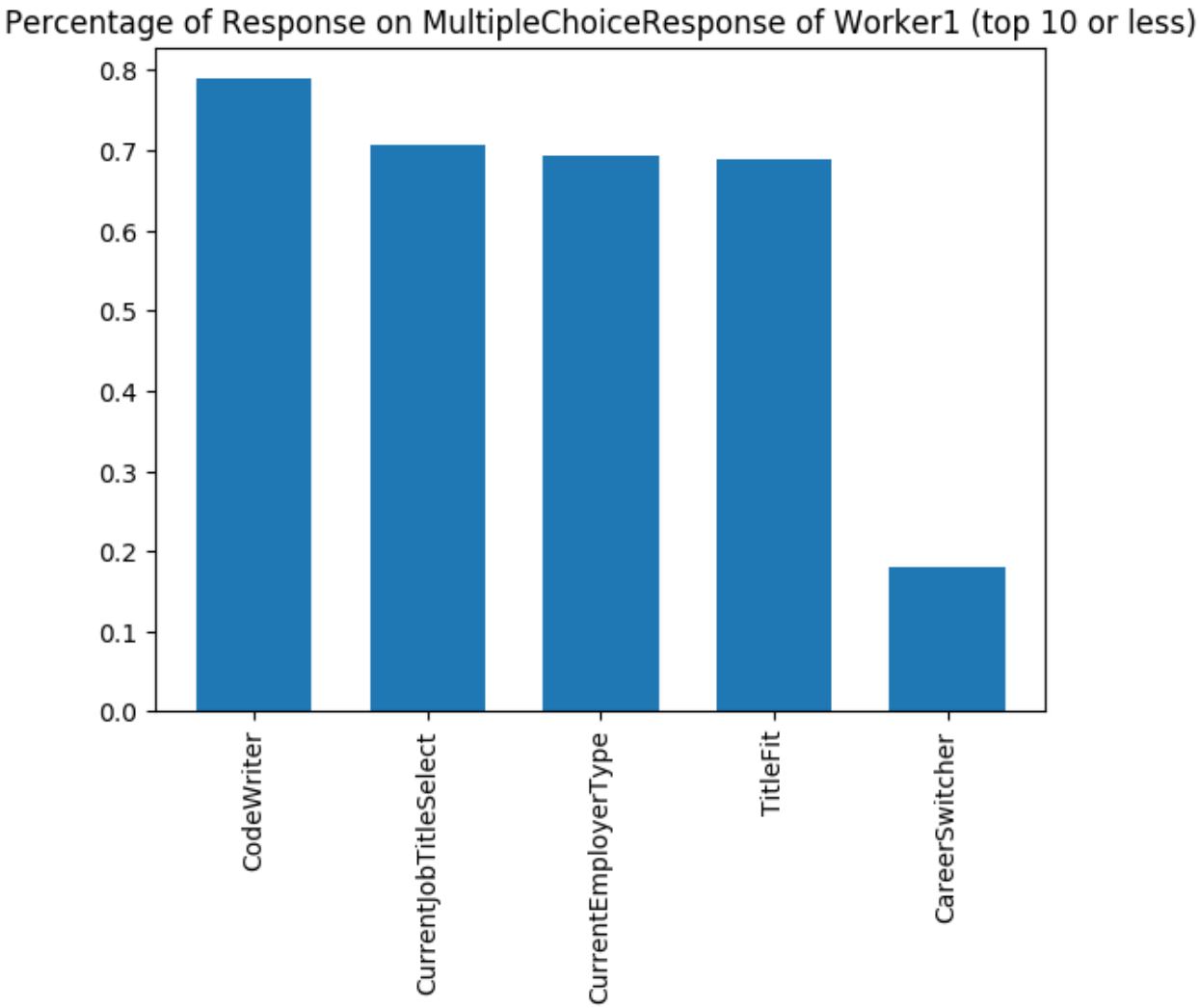
How important do you think the below skills or certifications are in getting a data science job? - Academic degree in related field (top 10 or less)



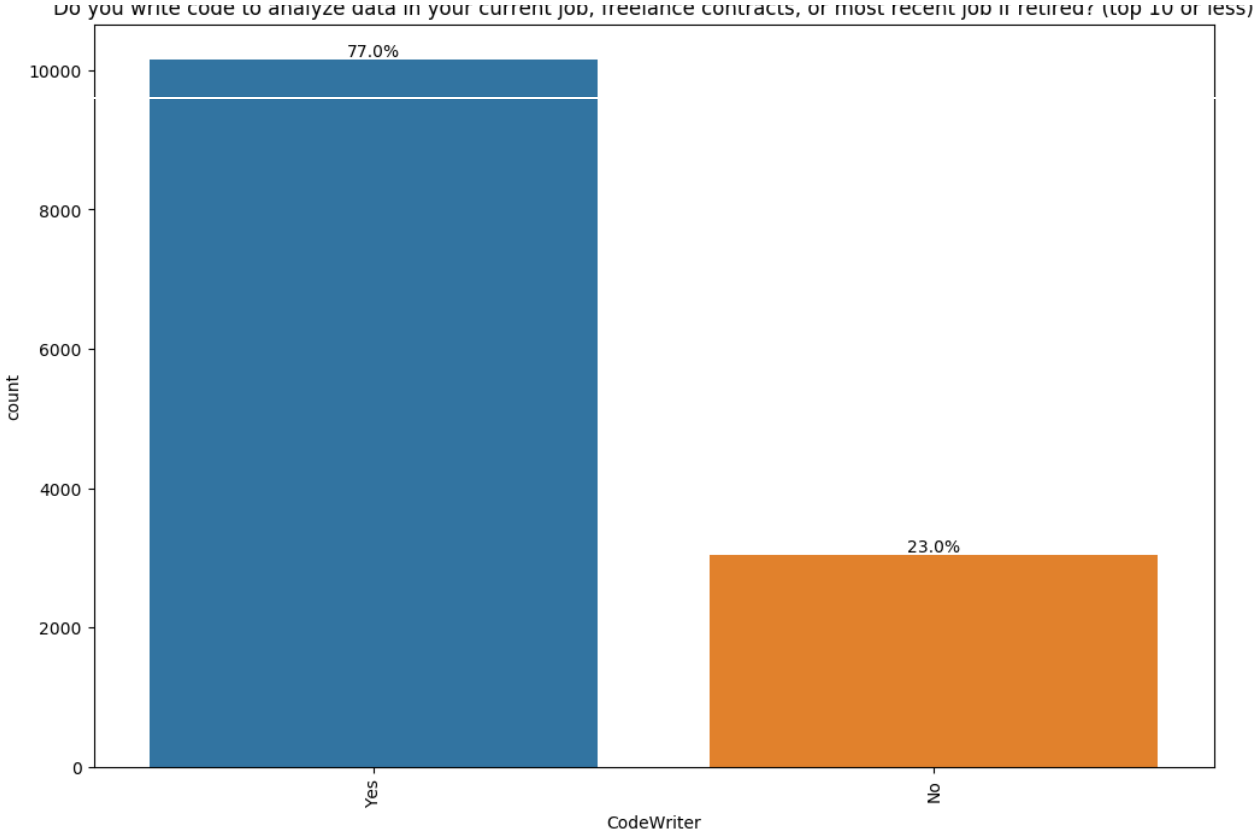
How important do you think the below skills or certifications are in getting a data science job? - Fluent in SQL (top 10 or less)

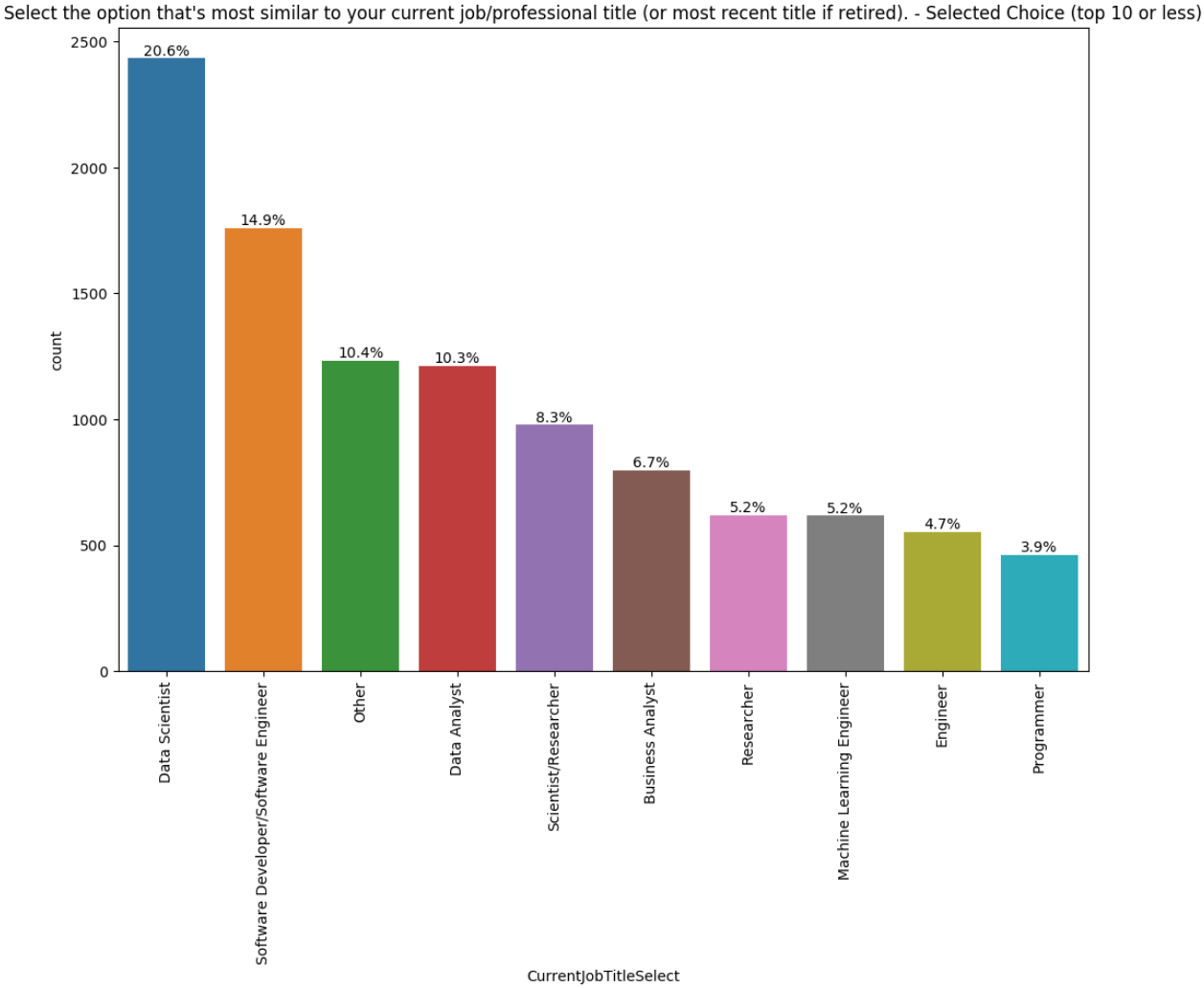


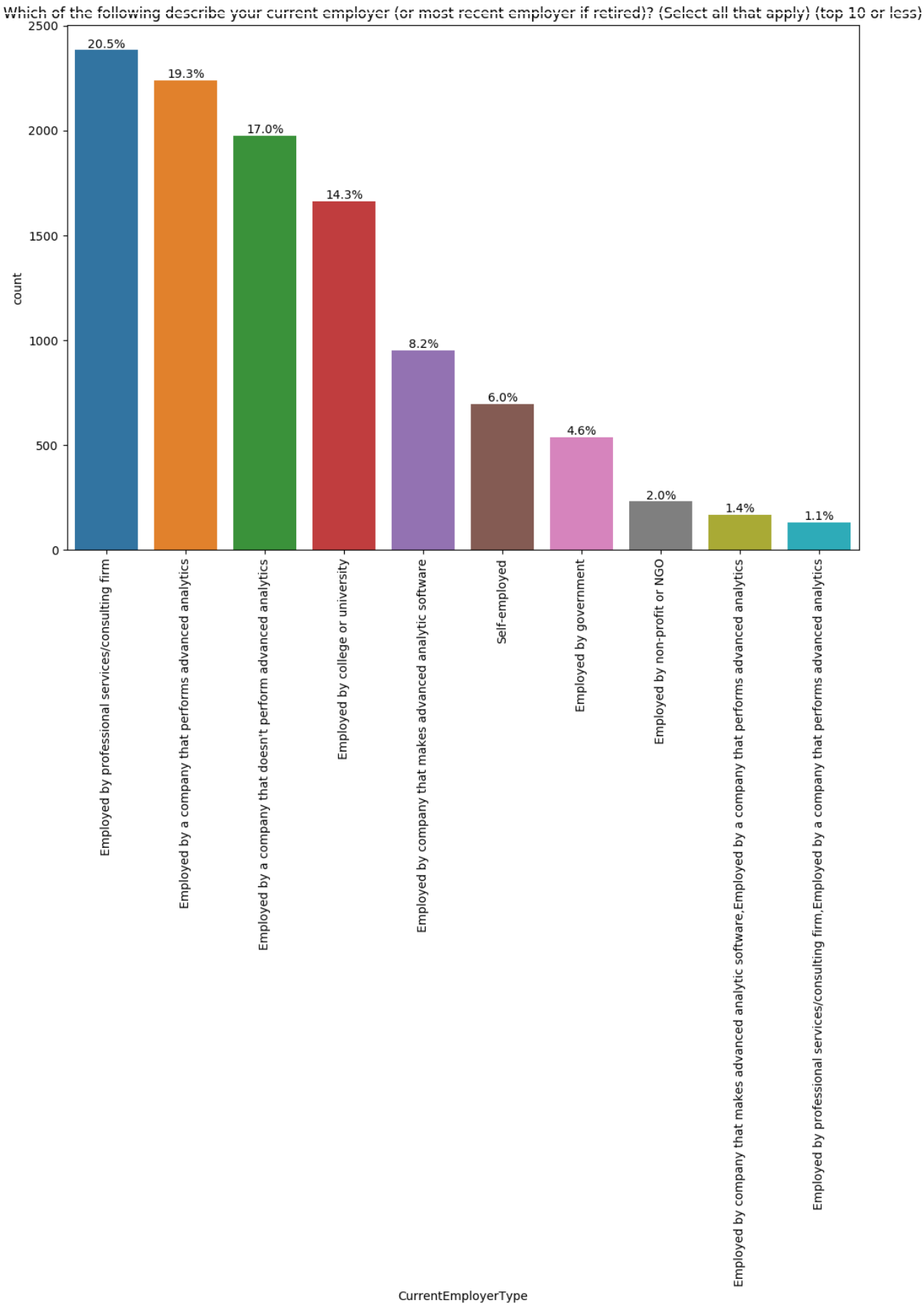




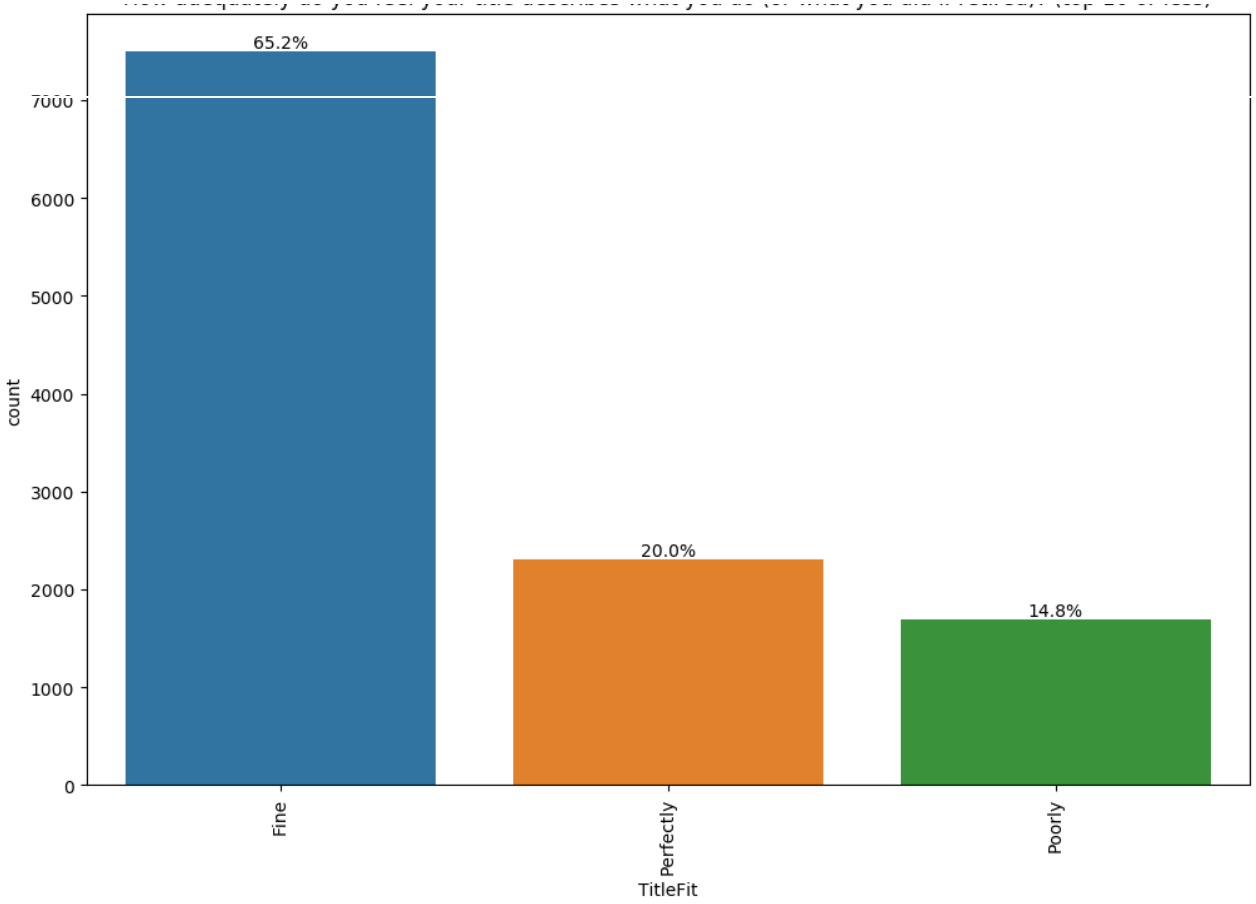
Do you write code to analyze data in your current job, freelance contracts, or most recent job if retired? (top 10 or less)

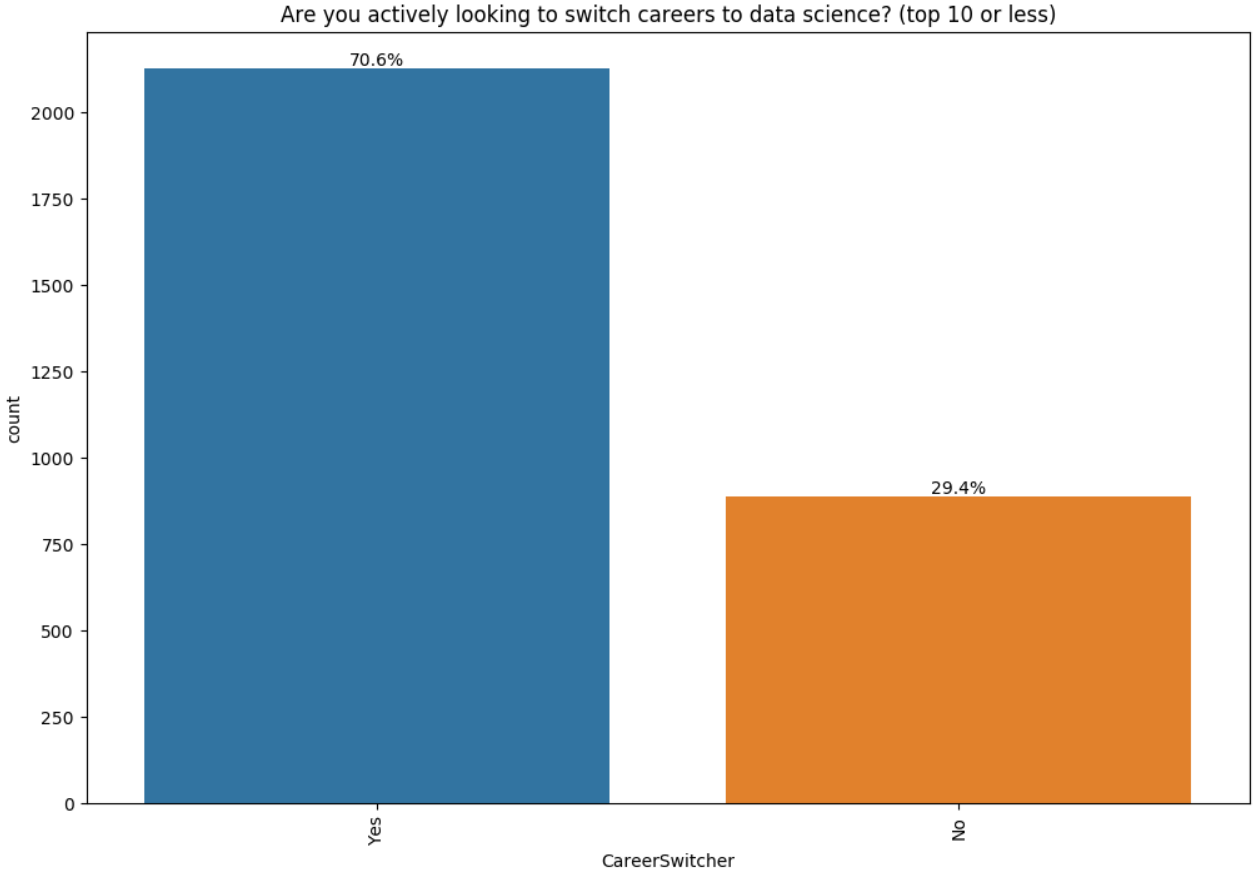

Do you write code to analyze data in your current job, freelance contracts, or most recent job if retired? (top 10 or less)
77.0%
23.0%
count
CodeWriter
Yes
No
0
2000
4000
6000
8000
10000

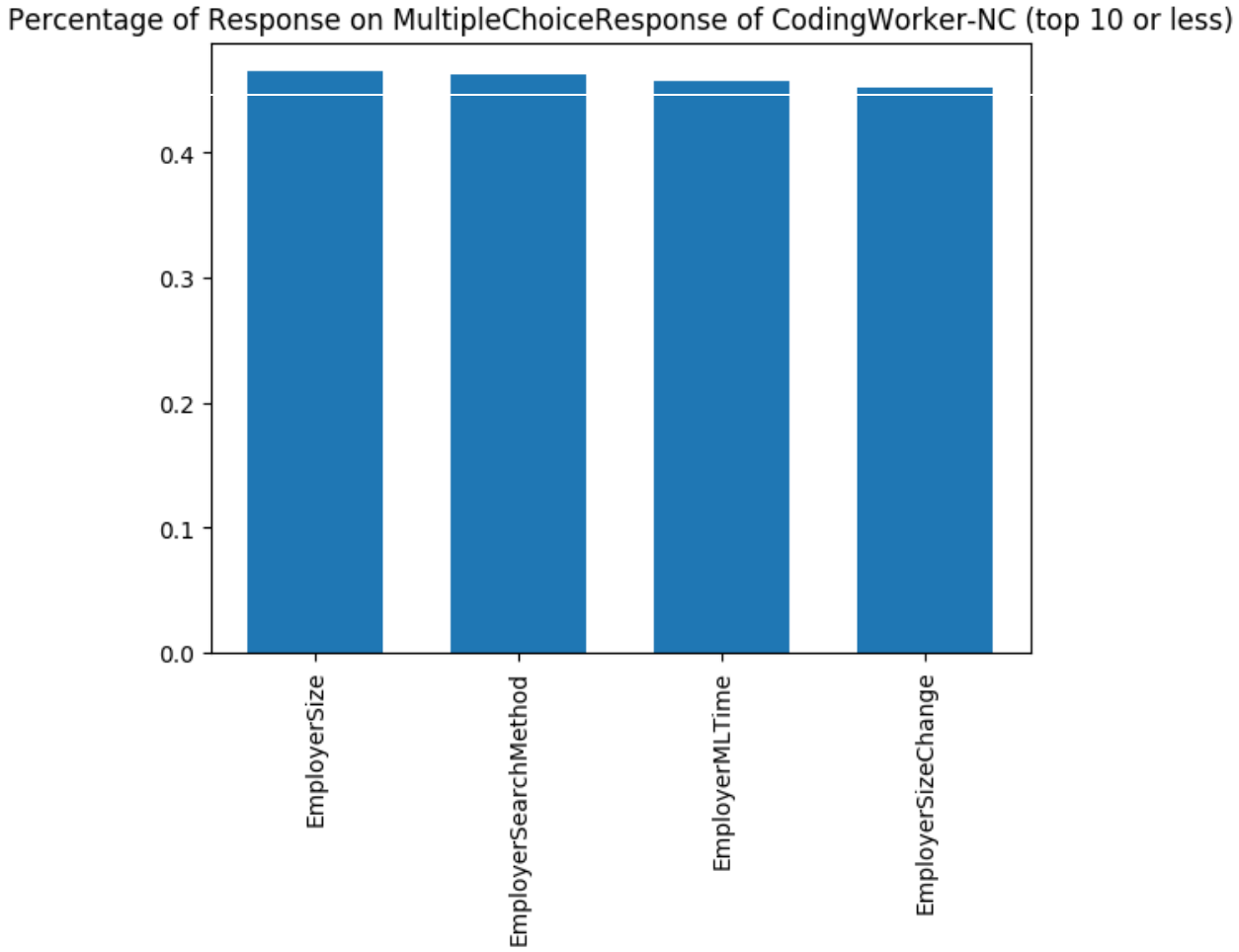

Select the option that's most similar to your current job/professional title (or most recent title if retired). - Selected Choice (top 10 or less)
20.6%
14.9%
10.4%
10.3%
8.3%
6.7%
5.2%
5.2%
4.7%
3.9%
count
CurrentJobTitleSelect
Data Scientist
Software Developer/Software Engineer
Other
Data Analyst
Scientist/Researcher
Business Analyst
Researcher
Machine Learning Engineer
Engineer
Programmer
0
500
1000
1500
2000
2500

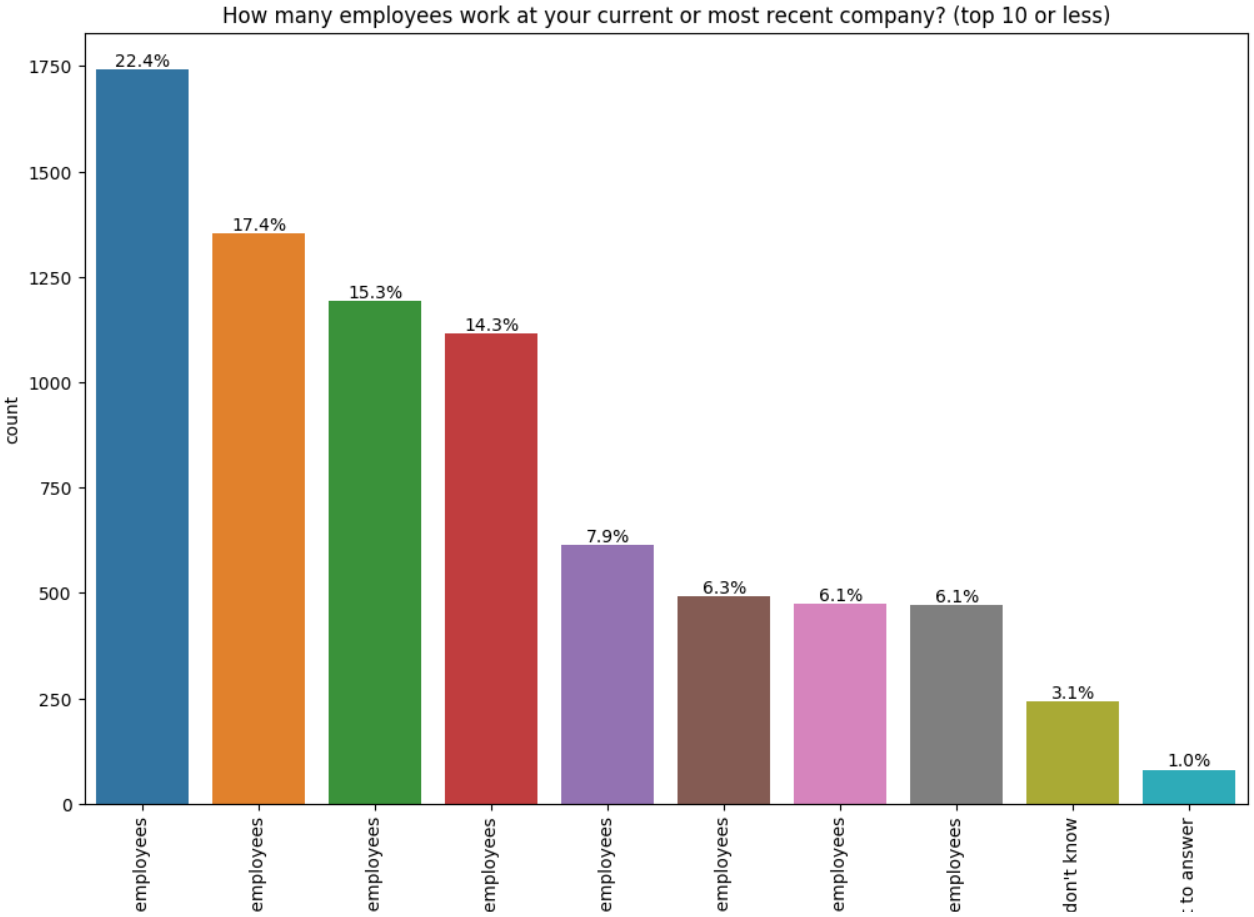


How adequately do you feel your title describes what you do (or what you did if retired)? (top 10 or less)

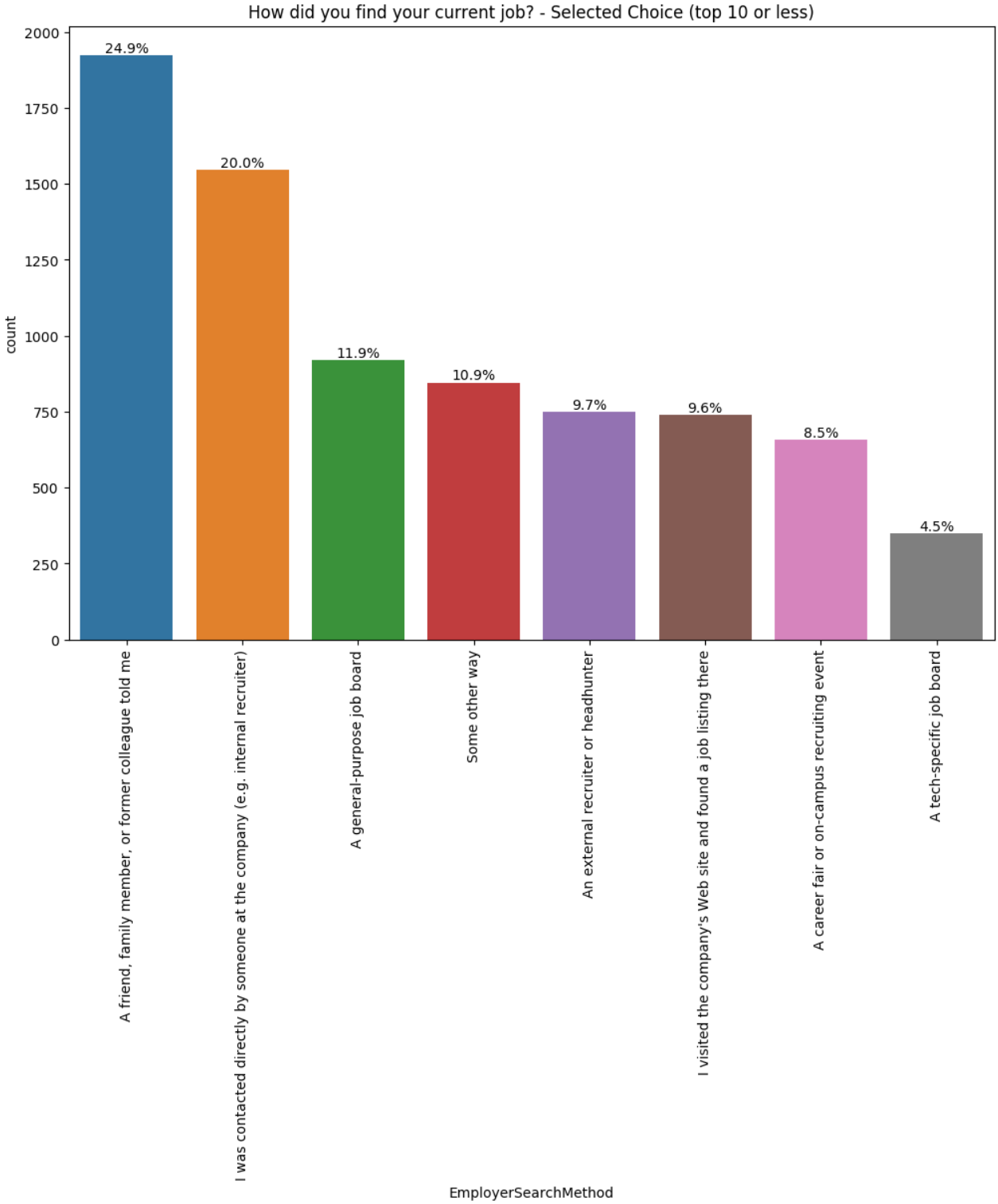

65.2%
20.0%
14.8%
count
TitleFit
Fine
Perfectly
Poorly

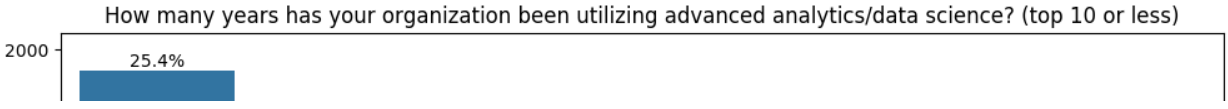

Are you actively looking to switch careers to data science? (top 10 or less)
70.6%
29.4%
count
CareerSwitcher
Yes
No

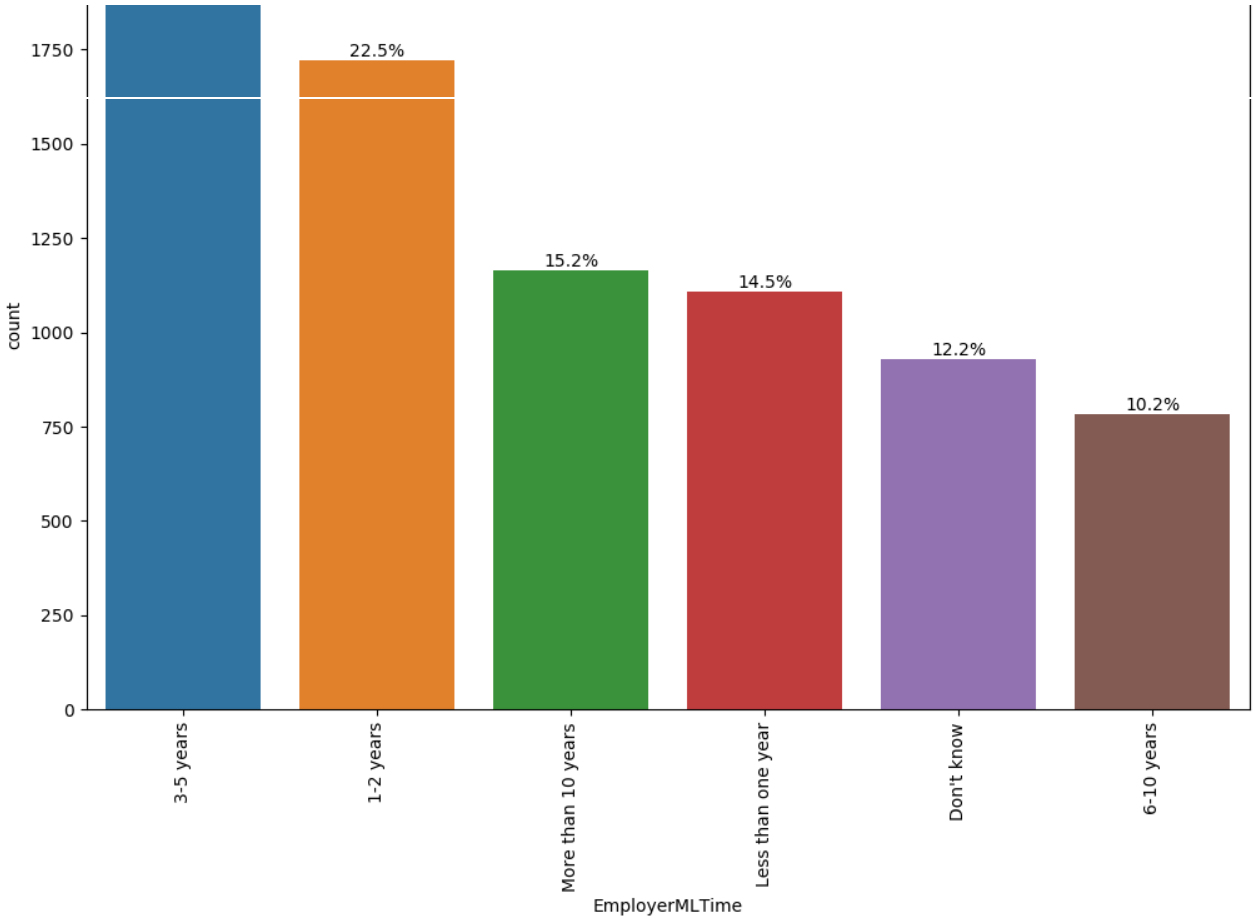

Percentage of Response on MultipleChoiceResponse of CodingWorker-NC (top 10 or less)
0.4
0.3
0.2
0.1
0.0
EmployerSize
EmployerSearchMethod
EmployerMLTime
EmployerSizeChange

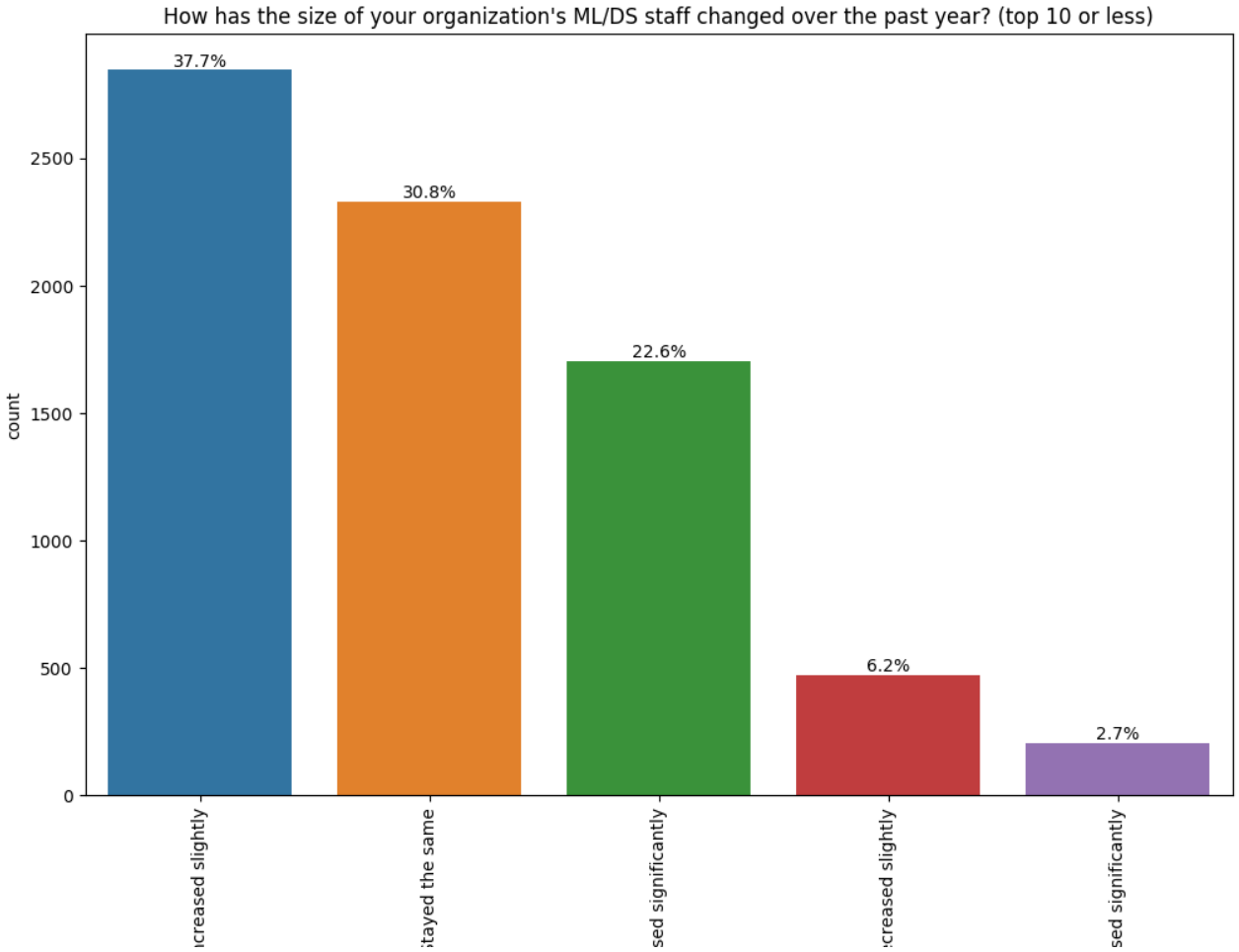

How many employees work at your current or most recent company? (top 10 or less)
count
1750
1500
1250
1000
750
500
250
0
22.4%
17.4%
15.3%
14.3%
7.9%
6.3%
6.1%
6.1%
3.1%
1.0%
employees
employees
employees
employees
employees
employees
employees
employees
don't know
: to answer


10,000 or more
100 to 499
20 to 99
1,000 to 4,999
500 to 999
10 to 19
Fewer than 10
5,000 to 9,999
I prefer not
EmployerSize

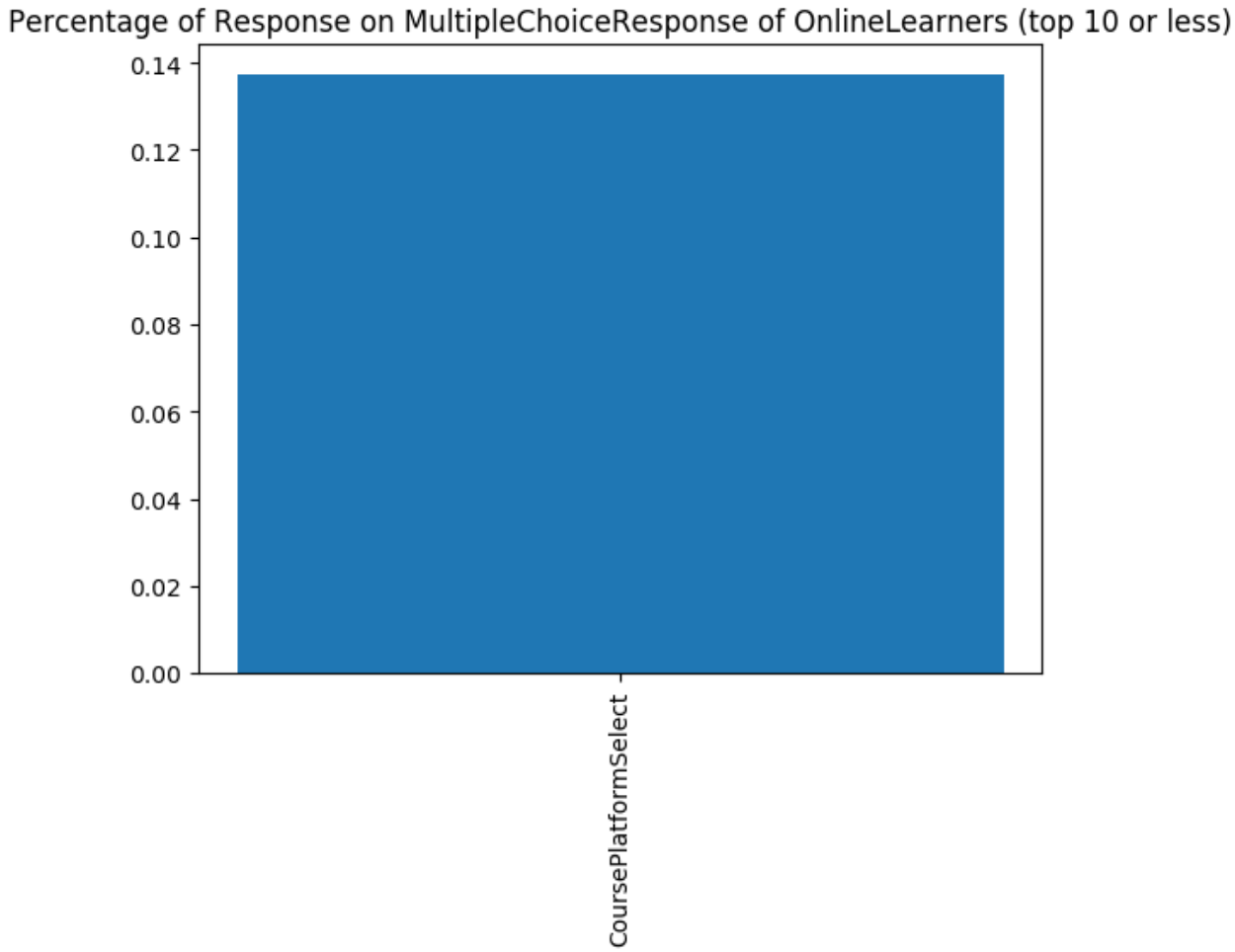

How did you find your current job? - Selected Choice (top 10 or less)
count
2000
1750
1500
1250
1000
750
500
250
0
24.9%
20.0%
11.9%
10.9%
9.7%
9.6%
8.5%
4.5%
A friend, family member, or former colleague told me
I was contacted directly by someone at the company (e.g. internal recruiter)
A general-purpose job board
Some other way
An external recruiter or headhunter
I visited the company's Web site and found a job listing there
A career fair or on-campus recruiting event
A tech-specific job board
EmployerSearchMethod

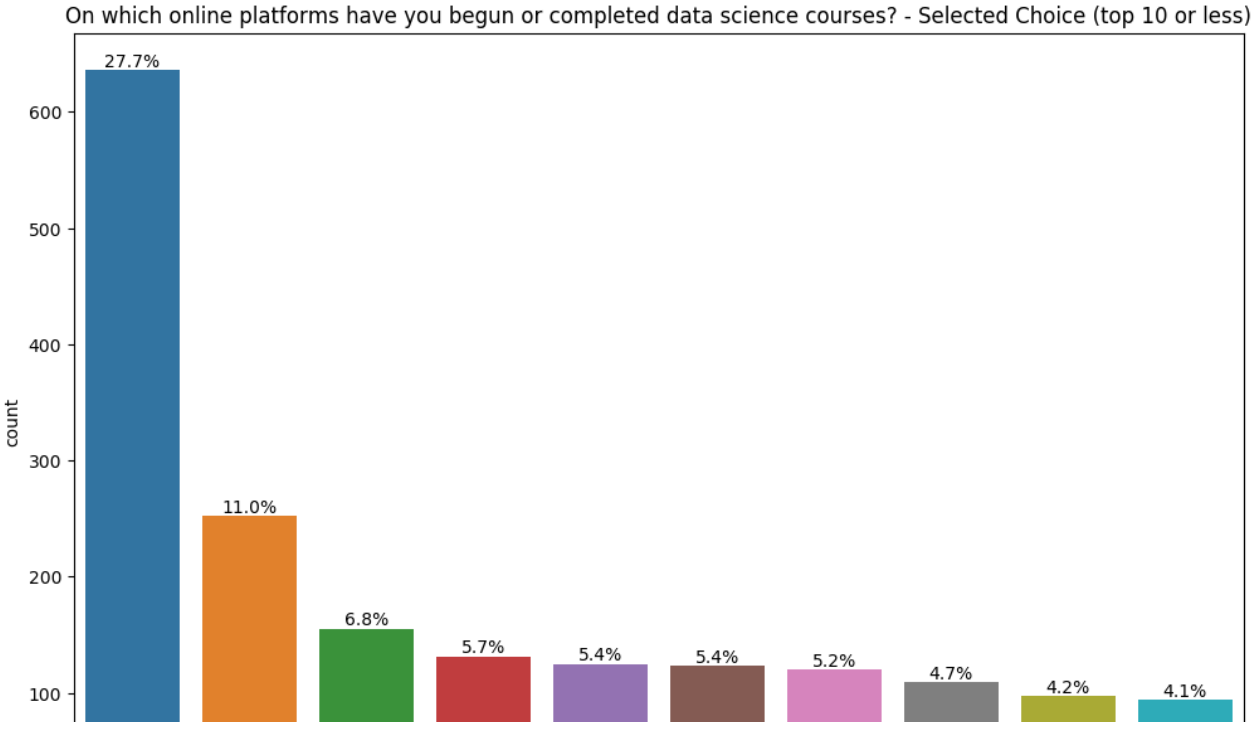

How many years has your organization been utilizing advanced analytics/data science? (top 10 or less)
2000
25.4%

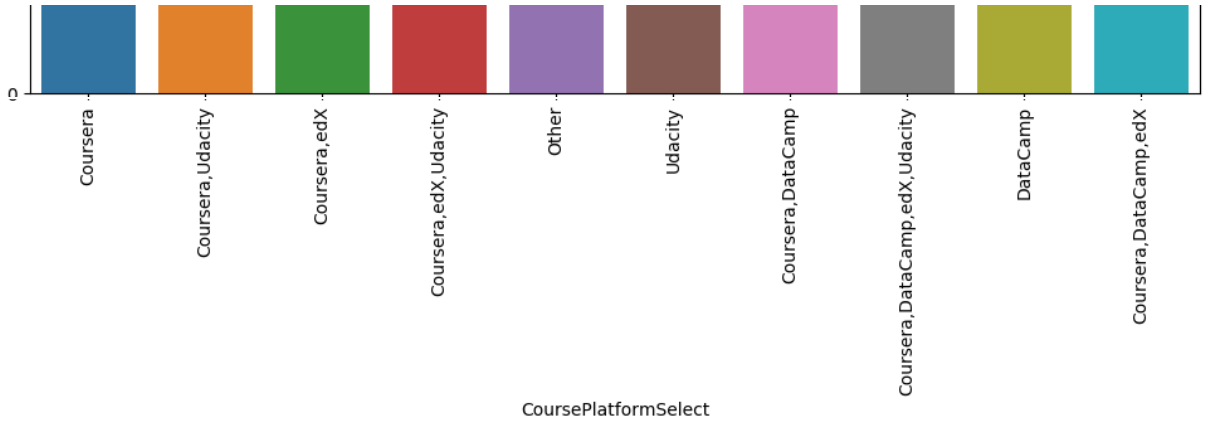

1750
1500
1250
1000
750
500
250
0
count
22.5%
15.2%
14.5%
12.2%
10.2%
3-5 years
1-2 years
More than 10 years
Less than one year
Don't know
6-10 years
EmployerMLTime

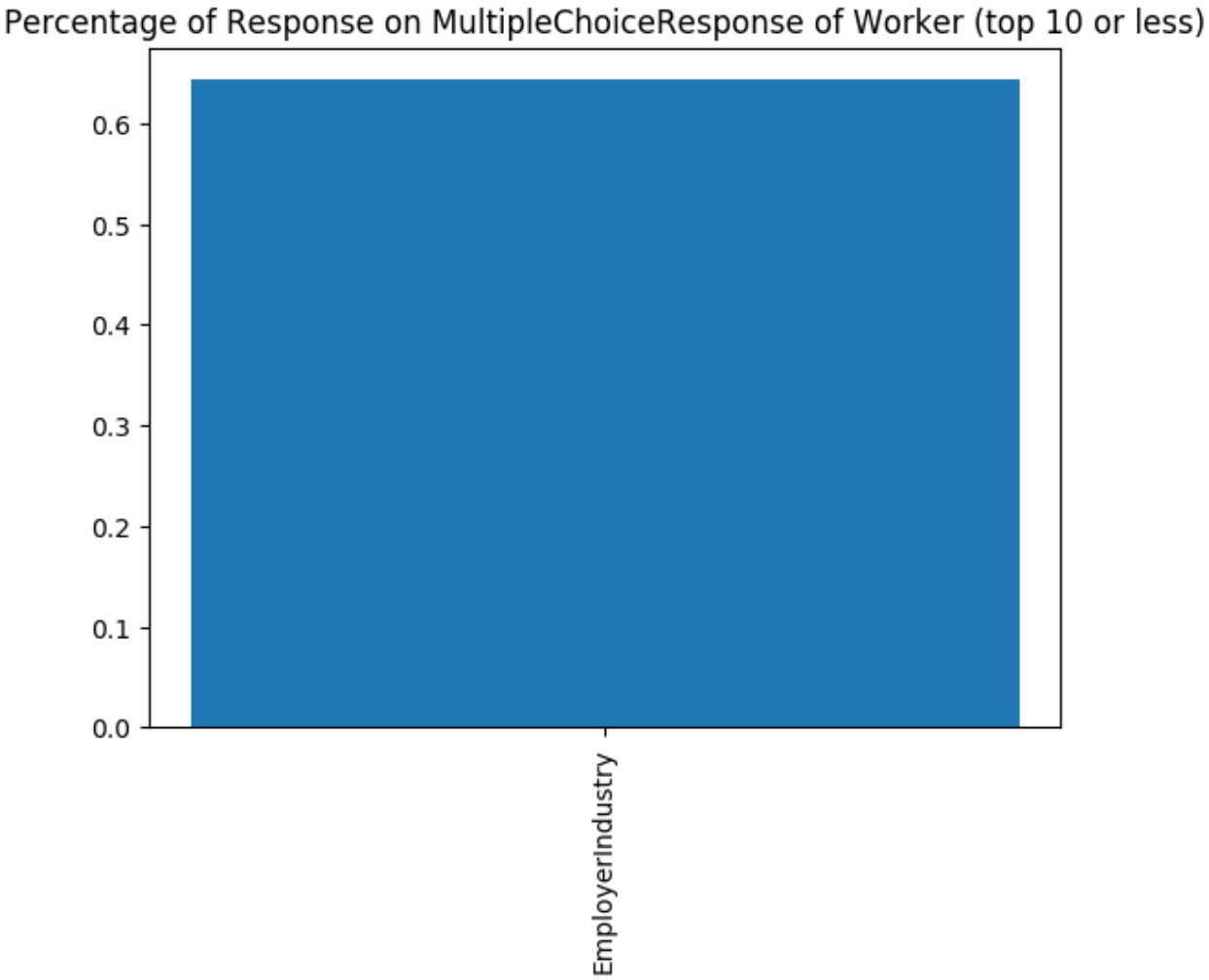

How has the size of your organization's ML/DS staff changed over the past year? (top 10 or less)
2500
2000
1500
1000
500
0
count
37.7%
30.8%
22.6%
6.2%
2.7%
ıcreased slightly
ıtayed the same
sed significantly
ıcreased slightly
sed significantly

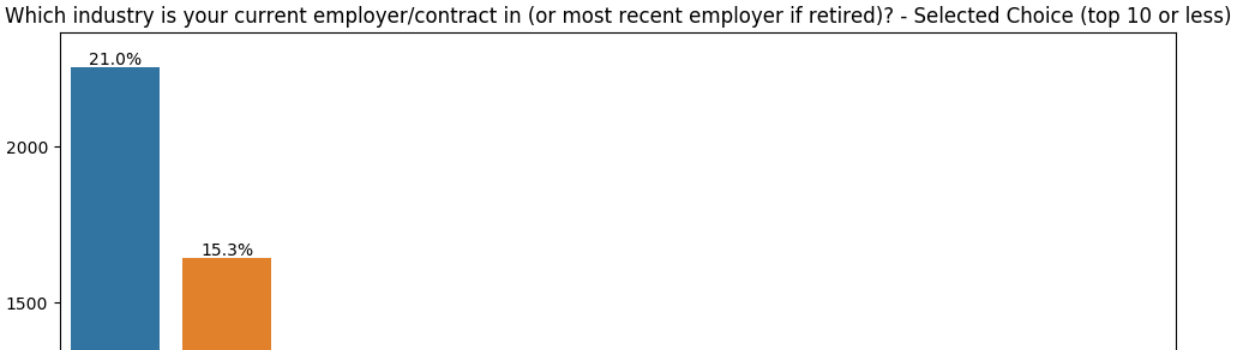

In
S
Increa
De
Decrea
EmployerSizeChange

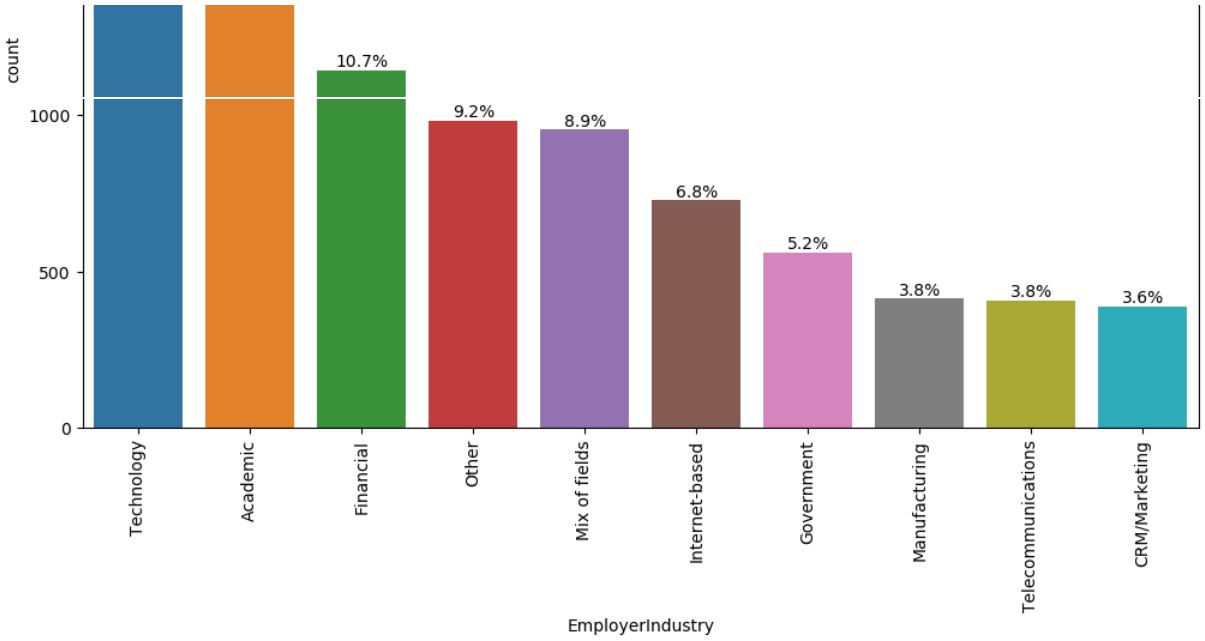

Percentage of Response on MultipleChoiceResponse of OnlineLearners (top 10 or less)
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
CoursePlatformSelect

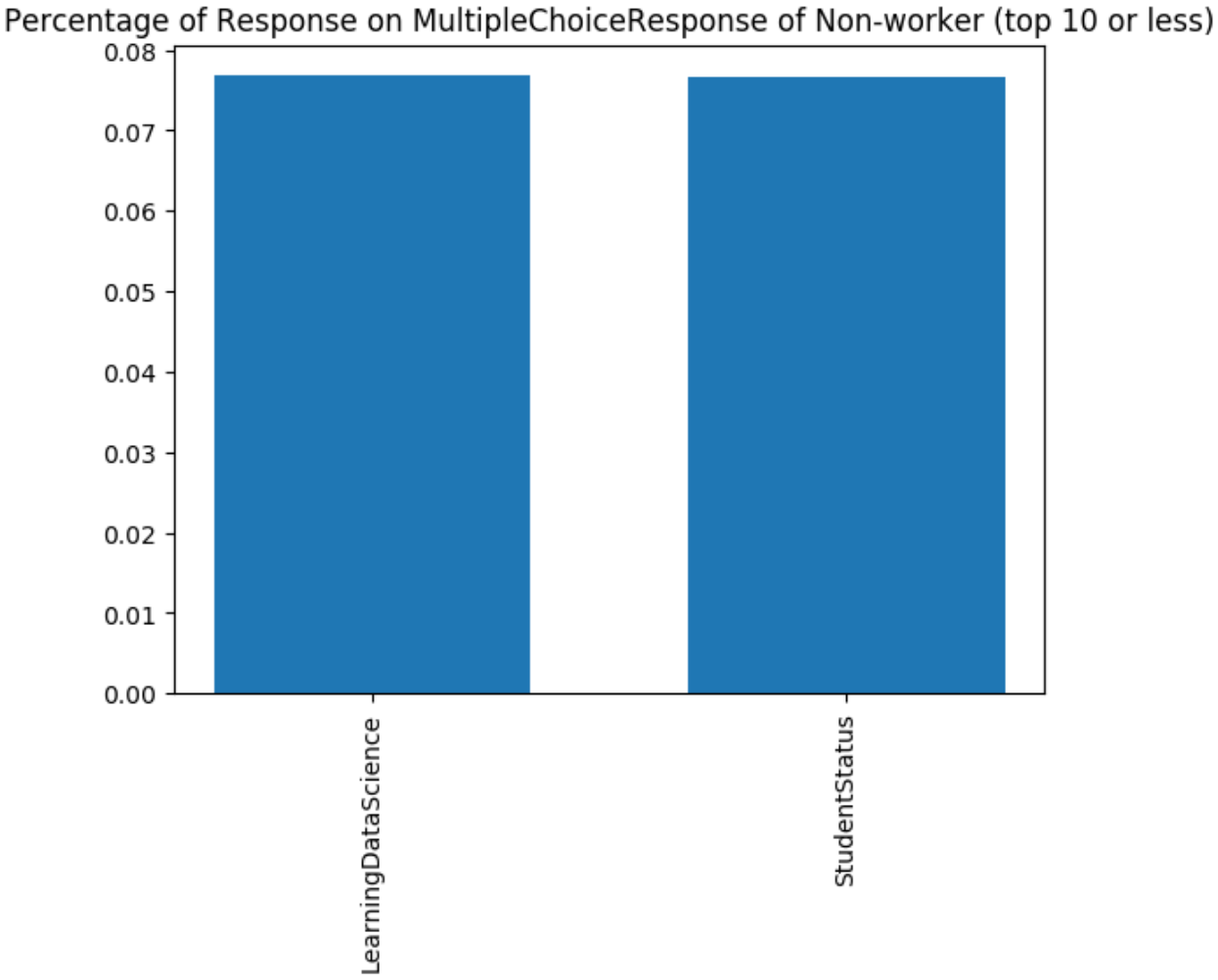

On which online platforms have you begun or completed data science courses? - Selected Choice (top 10 or less)
count
600
500
400
300
200
100
27.7%
11.0%
6.8%
5.7%
5.4%
5.4%
5.2%
4.7%
4.2%
4.1%

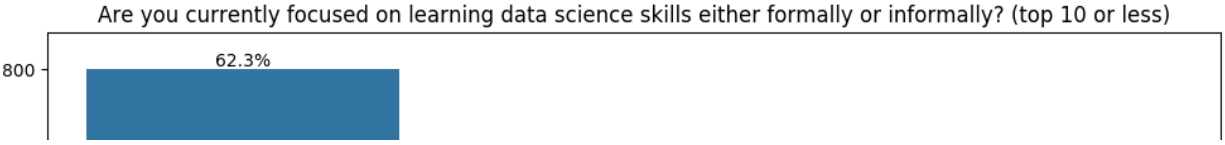

0
Coursera
Coursera,Udacity
Coursera,edX
Coursera,edX,Udacity
Other
Udacity
Coursera,DataCamp
Coursera,DataCamp,edX,Udacity
DataCamp
Coursera,DataCamp,edX
CoursePlatformSelect

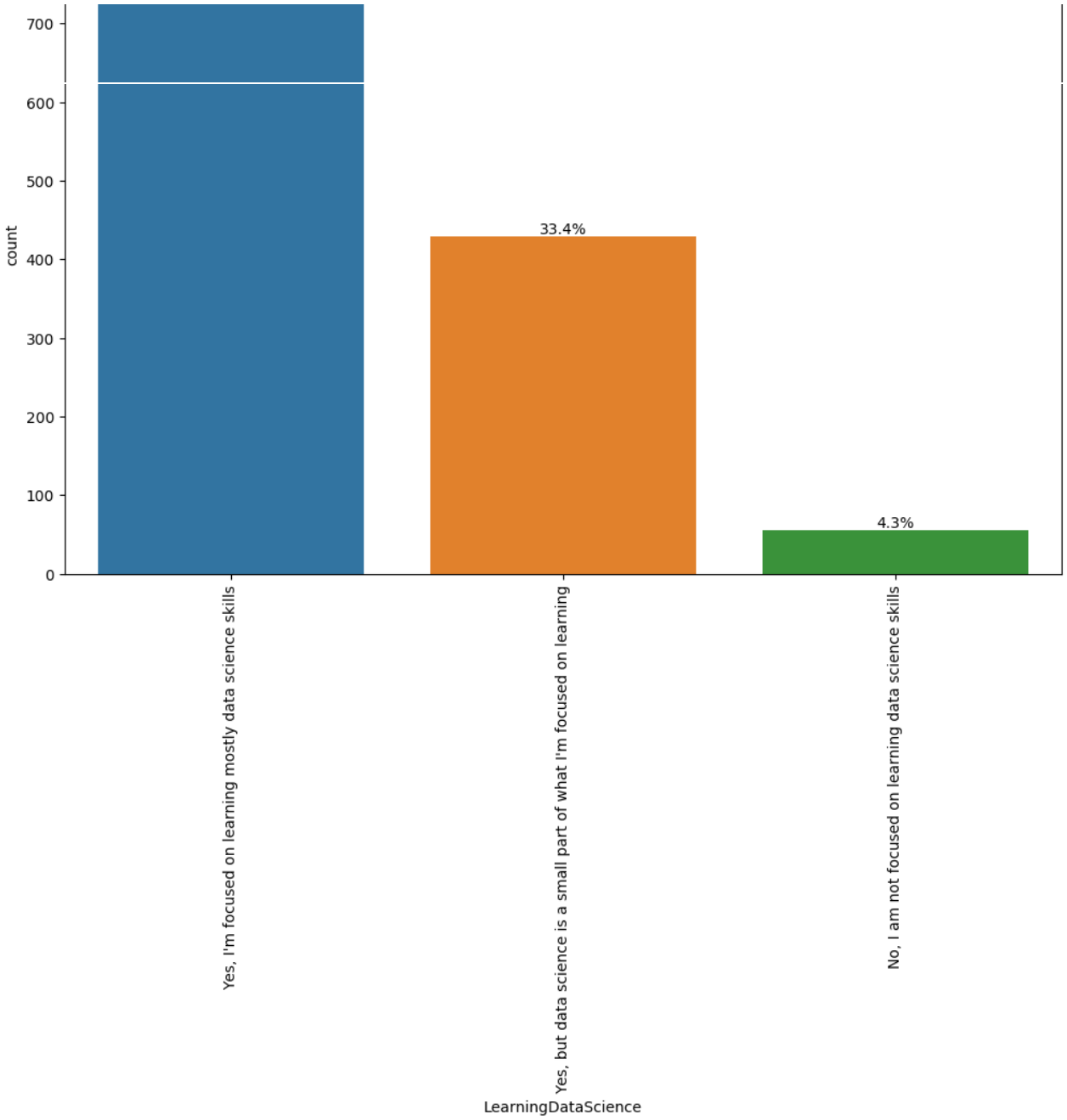

Percentage of Response on MultipleChoiceResponse of Worker (top 10 or less)
0.6
0.5
0.4
0.3
0.2
0.1
0.0
EmployerIndustry

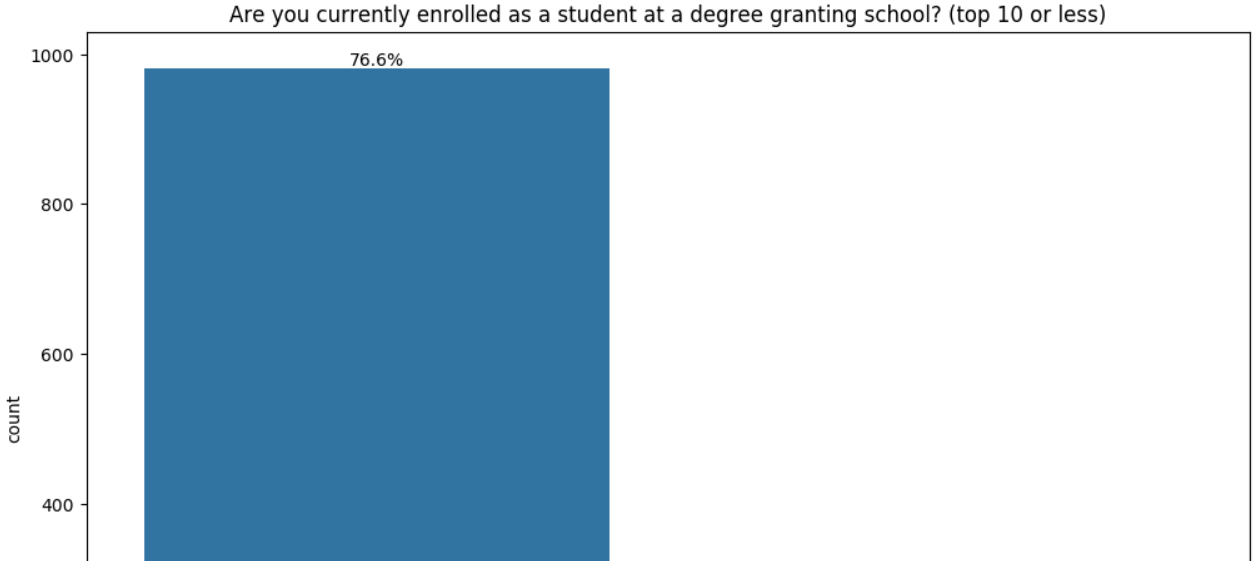

Which industry is your current employer/contract in (or most recent employer if retired)? - Selected Choice (top 10 or less)
21.0%
15.3%
2000
1500

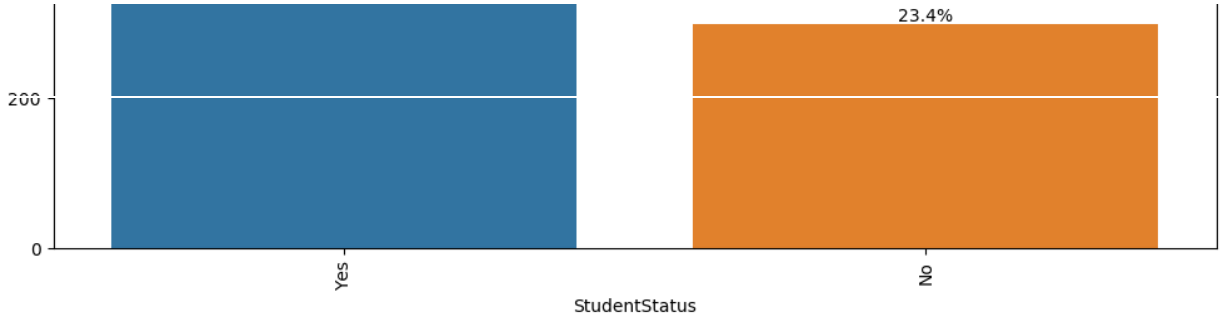

count
1000
500
0
10.7%
9.2%
8.9%
6.8%
5.2%
3.8%
3.8%
3.6%
Technology
Academic
Financial
Other
Mix of fields
Internet-based
Government
Manufacturing
Telecommunications
CRM/Marketing
EmployerIndustry

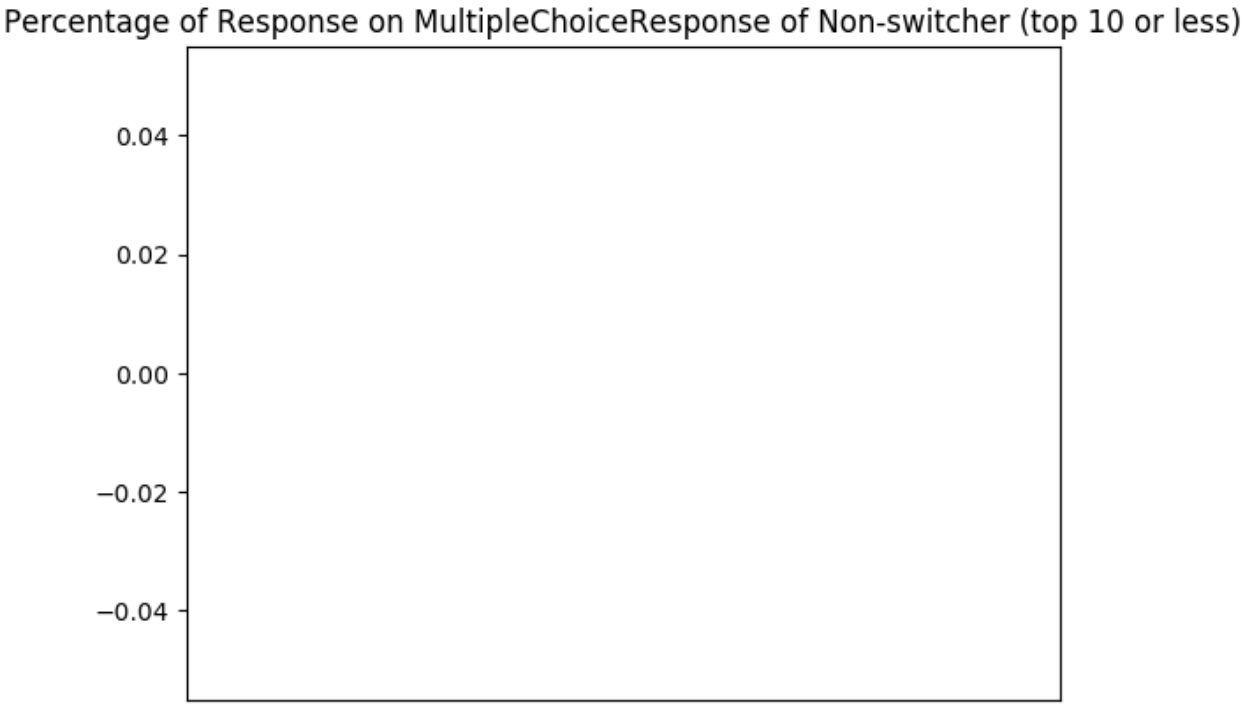

Percentage of Response on MultipleChoiceResponse of Non-worker (top 10 or less)
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
LearningDataScience
StudentStatus


Are you currently focused on learning data science skills either formally or informally? (top 10 or less)
800
62.3%


700
600
500
400
300
200
100
0
count
33.4%
4.3%
Yes, I'm focused on learning mostly data science skills
Yes, but data science is a small part of what I'm focused on learning
No, I am not focused on learning data science skills
LearningDataScience


Are you currently enrolled as a student at a degree granting school? (top 10 or less)
1000
800
600
400
count
76.6%


23.4%
200
0
Yes
No
StudentStatus


Percentage of Response on MultipleChoiceResponse of Non-switcher (top 10 or less)
0.04
0.02
0.00
−0.02
−0.04

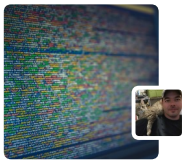
In [12]:

Comments (0)

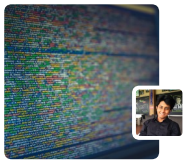
Sort by Select...

Click here to enter a comment...

Similar Kernels

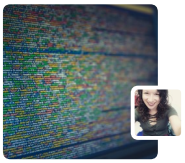


Analyzing the analyzers

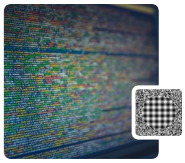


Data Science FAQ

An Interactive Deep Dive into Survey Results



Kaggle 2017 Survey Results



2017 State of Data Science - Kaggle survey