A

REPORT

ON

# Forecasting Millennial Career Trends

Submitted by,
**Summer Internship Project 2019**
**Shreyas Torgalmath**

SPOTLE LINK: https://spotle.ai/ShreyasTorgalmath

Click here to get CSV file: http://tiny.cc/rxag9y

# INDEX

In [1]:

```
import pandas as pd
%matplotlib inline
import matplotlib
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")
df=pd.read_csv("Data1.csv")
df["Serial"]=1
n=df["Serial"].sum()
print(n)
```

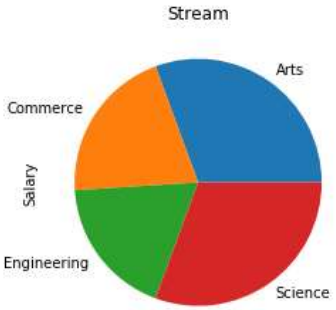
```
344
```

In [2]:
```
#What is the median salary that entry level employees expect grouped by stream?
df["Salary"]=df["expected salary"]
df["Salary"]=df["Salary"].replace("50,000 - 1,00,000",75000)
df["Salary"]=df["Salary"].replace("30,000 - 40,000",45000)
df["Salary"]=df["Salary"].replace("20,000 - 30,000",25000)
df["Salary"]=df["Salary"].replace("40,000 - 50,000",45000)
df["Salary"]=df["Salary"].replace("10,000 - 20,000",15000)
df["Salary"]=df["Salary"].replace("50,000-1,00,000",75000)
df["Salary"]=df["Salary"].replace("30,000-40,000",45000)
df["Salary"]=df["Salary"].replace("20,000-30,000",25000)
df["Salary"]=df["Salary"].replace("40,000-50,000",45000)
df["Salary"]=df["Salary"].replace("10,000-20,000",15000)
df["Salary"]=df["Salary"].replace("Less than 10,000",5000)
df["Salary"]=df["Salary"].replace("Greater than 1,00,000",150000)
df["Salary"]=df["Salary"].replace("other",50000)
df["Salary"]=df["Salary"].replace("According to my skills",50000)
df["Salary"]=df["Salary"].replace("50000-100000",75000)
df["Salary"]=df["Salary"].replace("<50000",25000)
df["Salary"]=df["Salary"].replace(">100000",150000)
df["Salary"]=df["Salary"].replace("Though salary is not a priority, it completely depends on the company",50000)
df["Salary"].fillna(50000, inplace=True)
Salary_Stream=df.groupby("Stream").Salary.median()
print(Salary_Stream)
```

```
Stream
Arts           75000.0
Commerce       50000.0
Engineering    45000.0
Science        75000.0
Name: Salary, dtype: float64
```

In [3]:
```
Salary_Stream.plot(kind='pie',title="Stream")
```

Out[3]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x29695a92c88>
```



In [4]:
```
#Preferred city to salary expectation for
City_Stream=df.groupby('work city').Salary.median()
City_Stream
```

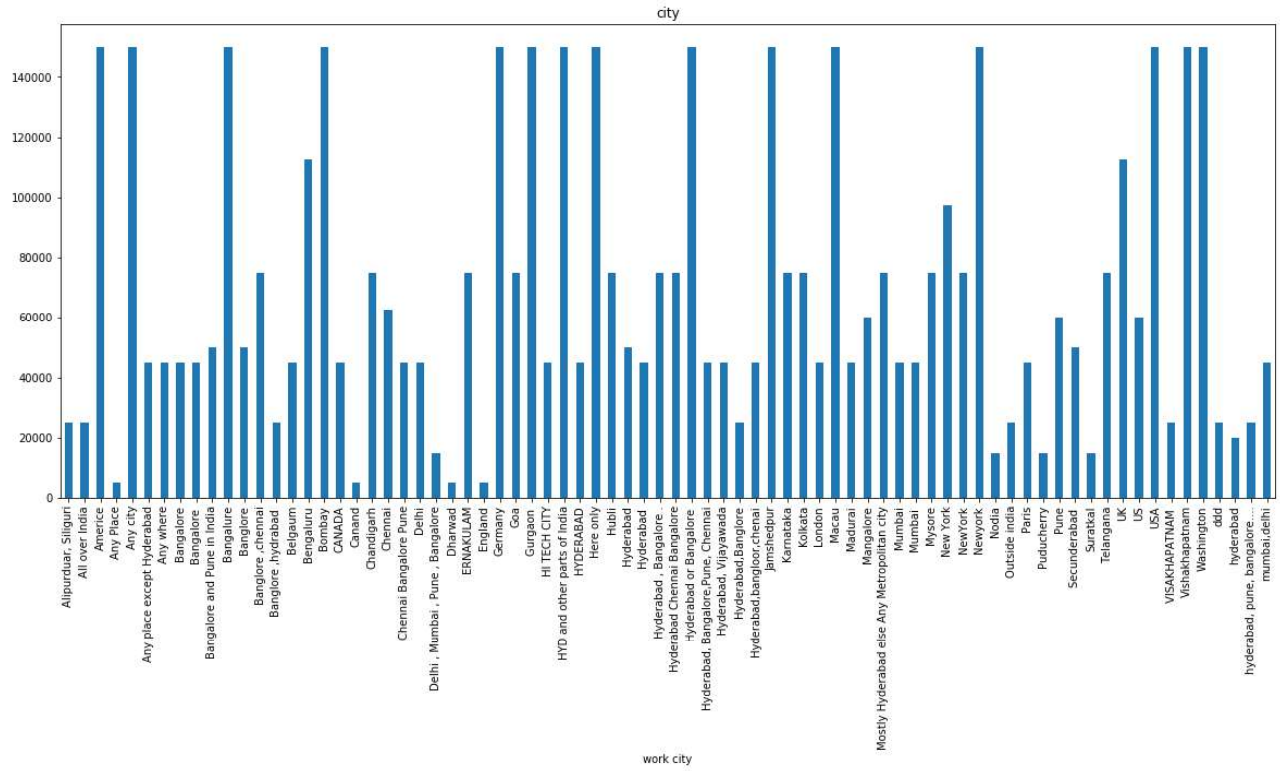
Out[4]:

| work city | |
|---|---|
| Alipurduar, Siliguri | 25000.0 |
| All over India | 25000.0 |
| Americe | 150000.0 |
| Any Place | 5000.0 |
| Any city | 150000.0 |
| Any place except Hyderabad | 45000.0 |
| Any where | 45000.0 |
| Bangalore | 45000.0 |
| Bangalore | 45000.0 |
| Bangalore and Pune in India | 50000.0 |
| Bangalure | 150000.0 |
| Banglore | 50000.0 |
| Banglore ,chennai | 75000.0 |
| Banglore ,hydrabad | 25000.0 |
| Belgaum | 45000.0 |
| Bengaluru | 112500.0 |
| Bombay | 150000.0 |
| CANADA | 45000.0 |
| Canand | 5000.0 |
| Chandigarh | 75000.0 |
| Chennai | 62500.0 |
| Chennai Bangalore Pune | 45000.0 |
| Delhi | 45000.0 |
| Delhi , Mumbai , Pune , Bangalore | 15000.0 |
| Dharwad | 5000.0 |
| ERNAKULAM | 75000.0 |
| England | 5000.0 |
| Germany | 150000.0 |
| Goa | 75000.0 |
| Gurgaon | 150000.0 |
| | ... |
| Kolkata | 75000.0 |
| London | 45000.0 |
| Macau | 150000.0 |
| Madurai | 45000.0 |
| Mangalore | 60000.0 |
| Mostly Hyderabad else Any Metropolitan city | 75000.0 |
| Mumbai | 45000.0 |
| Mumbai | 45000.0 |
| Mysore | 75000.0 |

```
New York                                    97500.0
NewYork                                     75000.0
Newyork                                    150000.0
Nodia                                       15000.0
Outside india                               25000.0
Paris                                       45000.0
Puducherry                                  15000.0
Pune                                        60000.0
Secunderabad                                50000.0
Suratkal                                    15000.0
Telangana                                   75000.0
UK                                         112500.0
US                                          60000.0
USA                                        150000.0
VISAKHAPATNAM                               25000.0
Vishakhapatnam                             150000.0
Washington                                 150000.0
ddd                                         25000.0
hyderabad                                   20000.0
hyderabad, pune, bangalore....              25000.0
mumbai,delhi                                45000.0
Name: Salary, Length: 76, dtype: float64
```

In [5]:
```
fig = plt.figure(figsize=(20, 8))
City_Stream.plot(kind='bar',title="city")
```

Out[5]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x296989be4e0>
```



In [6]:

```
#Preferred job industry by stream
df["Stream"].fillna('other', inplace = True)
Industry_Stream=df.groupby("Stream").sum()
df["stream"]=df["Stream"]
df["stream"]=df["stream"].replace("Arts",1)
df["stream"]=df["stream"].replace("Science",1000)
df["stream"]=df["stream"].replace("Commerce",1000000)
df["stream"]=df["stream"].replace("Engineering",1000000000)
df["stream"]=df["stream"].replace("other",1000000000000)
da=pd.DataFrame()
da["sum"]=df.groupby("Field to work in").stream.sum()
i=0
da["Engineering"]=0
da["Science"]=0
da["Commerce"]=0
da["Arts"]=0
da["Serial"]=1
da["other"]=0
nk=da["Serial"].sum()
print(nk)
del da["Serial"]
while i<nk:
    a=da["sum"][i]
    b=a%1000
    a=a/1000
    a=int(a)
    da["Arts"][i]=b
    b=a%1000
    a=int(a/1000)
    da["Science"][i]=b
    b=a%1000
    a=int(a/1000)
    da["Engineering"][i]=b
    b=a%1000
    a=int(a/1000)
    da["Commerce"][i]=b
    b=a%1000
    a=int(a/1000)
    da["other"][i]=b
    i=i+1
```

```
del da["sum"]
da
```

```
15
```

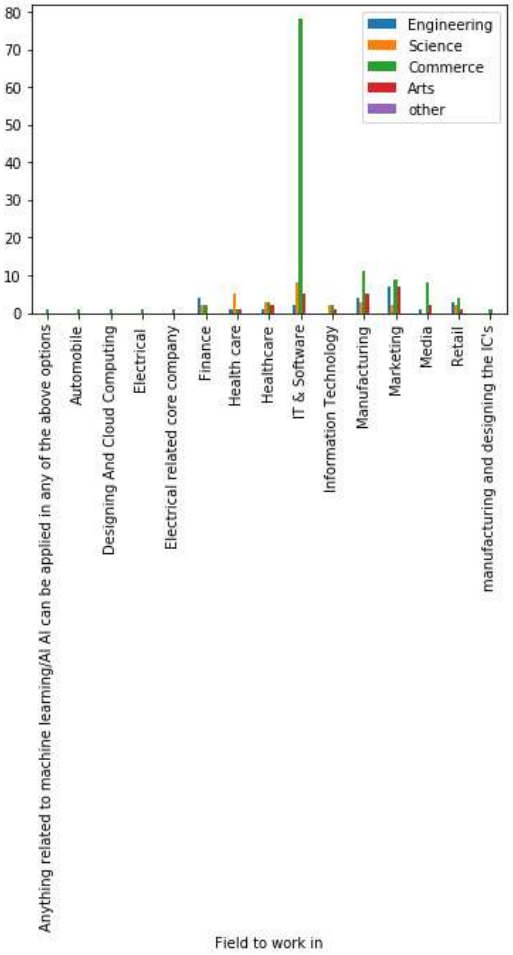
Out[6]:

| Field to work in | Engineering | Science | Commerce | Arts | other |
|---|---|---|---|---|---|
| Anything related to machine learning/AI AI can be applied in any of the above options | 0 | 0 | 1 | 0 | 0 |
| Automobile | 0 | 0 | 1 | 0 | 0 |
| Designing And Cloud Computing | 0 | 0 | 1 | 0 | 0 |
| Electrical | 0 | 0 | 1 | 0 | 0 |
| Electrical related core company | 0 | 0 | 1 | 0 | 0 |
| Finance | 4 | 2 | 2 | 0 | 0 |
| Health care | 1 | 5 | 1 | 1 | 0 |
| Healthcare | 1 | 3 | 3 | 2 | 0 |
| IT & Software | 2 | 8 | 78 | 5 | 0 |
| Information Technology | 0 | 2 | 2 | 1 | 0 |
| Manufacturing | 4 | 3 | 11 | 5 | 0 |
| Marketing | 7 | 2 | 9 | 7 | 0 |
| Media | 1 | 0 | 8 | 2 | 0 |
| Retail | 3 | 2 | 4 | 1 | 0 |
| manufacturing and designing the IC's | 0 | 0 | 1 | 0 | 0 |

In [7]:
```
fig = plt.figure(figsize=(15, 5))
da.plot(kind='bar')
```

Out[7]: <matplotlib.axes._subplots.AxesSubplot at 0x29698f12518>

<Figure size 1080x360 with 0 Axes>


Engineering
Science
Commerce
Arts
other
80
70
60
50
40
30
20
10
0
Anything related to machine learning/AI AI can be applied in any of the above options
Automobile
Designing And Cloud Computing
Electrical
Electrical related core company
Finance
Health care
Healthcare
IT & Software
Information Technology
Manufacturing
Marketing
Media
Retail
manufacturing and designing the IC's
Field to work in

In [8]:
```
#Preferred salary by college tier
College_Stream=df.groupby("College").Salary.median()
College_Stream
```

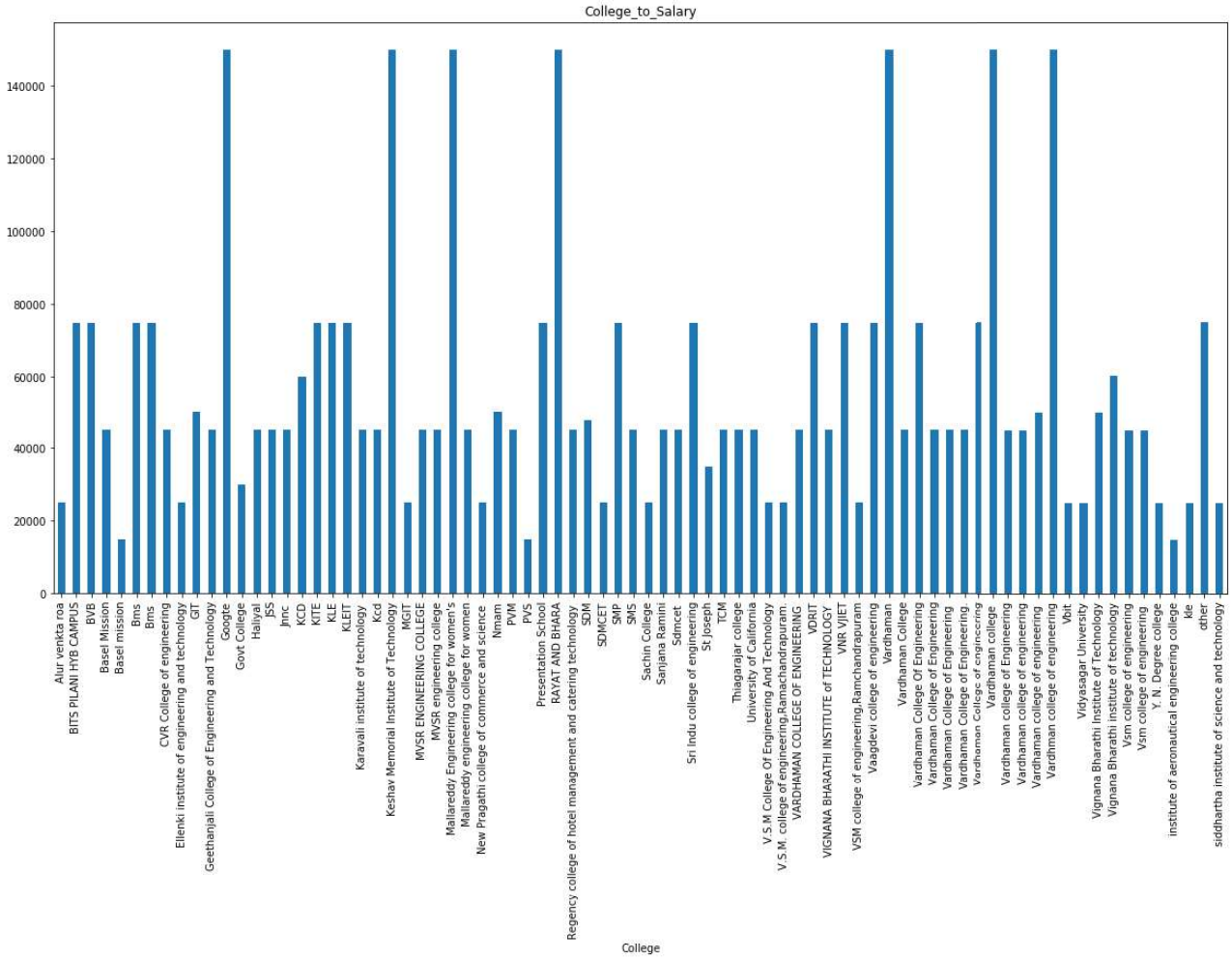
Out[8]:

| College | |
|---|---|
| Alur venkta roa | 25000.0 |
| BITS PILANI HYB CAMPUS | 75000.0 |
| BVB | 75000.0 |
| Basel Mission | 45000.0 |
| Basel mission | 15000.0 |
| Bms | 75000.0 |
| Bms | 75000.0 |
| CVR College of engineering | 45000.0 |
| Ellenki institute of engineering and technology | 25000.0 |
| GIT | 50000.0 |
| Geethanjali College of Engineering and Technology | 45000.0 |
| Googte | 150000.0 |
| Govt College | 30000.0 |
| Haliyal | 45000.0 |
| JSS | 45000.0 |
| Jnnc | 45000.0 |
| KCD | 60000.0 |
| KITE | 75000.0 |
| KLE | 75000.0 |
| KLEIT | 75000.0 |
| Karavali institute of technology | 45000.0 |
| Kcd | 45000.0 |
| Keshav Memorial Institute of Technology | 150000.0 |
| MGIT | 25000.0 |
| MVSR ENGINEERING COLLEGE | 45000.0 |
| MVSR engineering college | 45000.0 |
| Mallareddy Engineering college for women's | 150000.0 |
| Mallareddy engineering college for women | 45000.0 |
| New Pragathi college of commerce and science | 25000.0 |
| Nmam | 50000.0 |
| | ... |
| V.S.M. college of engineering,Ramachandrapuram. | 25000.0 |
| VARDHAMAN COLLEGE OF ENGINEERING | 45000.0 |
| VDRIT | 75000.0 |
| VIGNANA BHARATHI INSTITUTE of TECHNOLOGY | 45000.0 |
| VNR VJIET | 75000.0 |
| VSM college of engineering,Ramchandrapuram | 25000.0 |
| Vaagdevi college of engineering | 75000.0 |
| Vardhaman | 150000.0 |
| Vardhaman College | 45000.0 |

```
Vardhaman College Of Engineering                    75000.0
Vardhaman College of Engineering                    45000.0
Vardhaman College of Engineering                    45000.0
Vardhaman College of Engineering.                   45000.0
Vardhaman College of engineering                    75000.0
Vardhaman college                                  150000.0
Vardhaman college of Engineering                    45000.0
Vardhaman college of engineering                    45000.0
Vardhaman college of engineering                    50000.0
Vardhman college of engineering                    150000.0
Vbit                                                25000.0
Vidyasagar University                               25000.0
Vignana Bharathi Institute of Technology            50000.0
Vignana Bharathi institute of technology            60000.0
Vsm college of engineering                          45000.0
Vsm college of engineering                          45000.0
Y. N. Degree college                                25000.0
institute of aeronautical engineering college       15000.0
kle                                                 25000.0
other                                               75000.0
siddhartha institute of science and technology      25000.0
Name: Salary, Length: 78, dtype: float64
```

In [9]:
```
fig = plt.figure(figsize=(20, 10))
College_Stream.plot(kind='bar',title="College_to_Salary")
```

Out[9]: <matplotlib.axes._subplots.AxesSubplot at 0x29698ec1dd8>

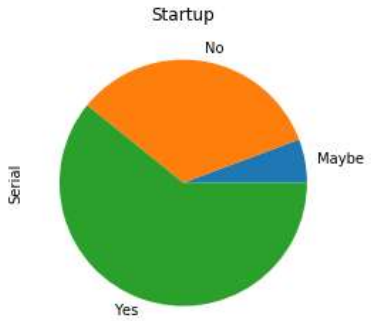

College_to_Salary
140000
120000
100000
80000
60000
40000
20000
0
Alur venkta roa
BITS PILANI HYB CAMPUS
BVB
Basel Mission
Basel mission
Bms
Bms
CVR College of engineering
Ellenki institute of engineering and technology
GIT
Geethanjali College of Engineering and Technology
Googte
Govt College
Haliyal
JSS
Jnnc
KCD
KTE
KLE
KLEIT
Karavali institute of technology
Kcd
Keshav Memorial Institute of Technology
MGIT
MVSR ENGINEERING COLLEGE
MVSR engineering college
Mallareddy Engineering college for women's
Mallareddy engineering college for women
New Pragathi college of commerce and science
Nmam
PVM
PVS
Presentation School
RAYAT AND BHARA
Regency college of hotel management and catering technology
SDM
SDMCET
SNP
SNS
Sachin College
Sanjana Ramini
Sdmcet
Sri Indu college of engineering
St Joseph
TCM
Thiagarajar college
University of California
V.S.M College Of Engineering And Technology
V.S.M. college of engineering,Ramachandrapuram.
VARDHAMAN COLLEGE OF ENGINEERING
VDRIT
VIGNANA BHARATHI INSTITUTE of TECHNOLOGY
VNR VJIET
VSM college of engineering,Ramchandrapuram
Vaagdevi college of engineering
Vardhaman
Vardhaman College
Vardhaman College Of Engineering
Vardhaman College of Engineering
Vardhaman College of Engineering
Vardhaman College of engineering
Vardhaman college
Vardhaman college of Engineering
Vardhaman college of engineering
Vardhaman college of engineering
Vbit
Vidyasagar University
Vignana Bharathi Institute of Technology
Vignana Bharathi institute of technology
Vsm college of engineering
Vsm college of engineering
Y. N. Degree college
institute of aeronautical engineering college
kle
other
siddhartha institute of science and technology
College

In [10]:
```
#Percentage of permanent job seekers vs those interested in entrepreneurship
df["Serial"]=1
df[' creating a startup']=df[' creating a startup'].replace("May be","Maybe")
Job=df.groupby(' creating a startup').Serial.sum()
Permanent_Job=Job.No+(Job.Maybe)/2
Entrepreneurship=Job.Yes+(Job.Maybe)/2
Total_Job=Job.Yes+Job.Maybe+Job.No
Percentage_of_permeanent_job=Permanent_Job*100/Total_Job
print(Percentage_of_permeanent_job)
```

```
36.295180722891565
```

In [11]:
```
Job.plot(kind='pie',title="Startup")
```

Out[11]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x296997eeb70>
```



In [12]:

```
# Which city shows more tilt to entrepreneurship vs fixed job
i=0
df["City_preference"]=0
while i<n:
    if(df[' creating a startup'][i]=="Yes"):
        df["City_preference"][i]=1
    elif(df[' creating a startup'][i]=="No"):
        df["City_preference"][i]=-1
    i=i+1
```

In [13]:
```
City_more_preference=df.groupby('City').City_preference.sum()
print(City_more_preference)
dj=pd.DataFrame()
dj["City_more_preference"]=City_more_preference
marker=City_more_preference.max()
k=dj[dj.City_more_preference==marker]
```

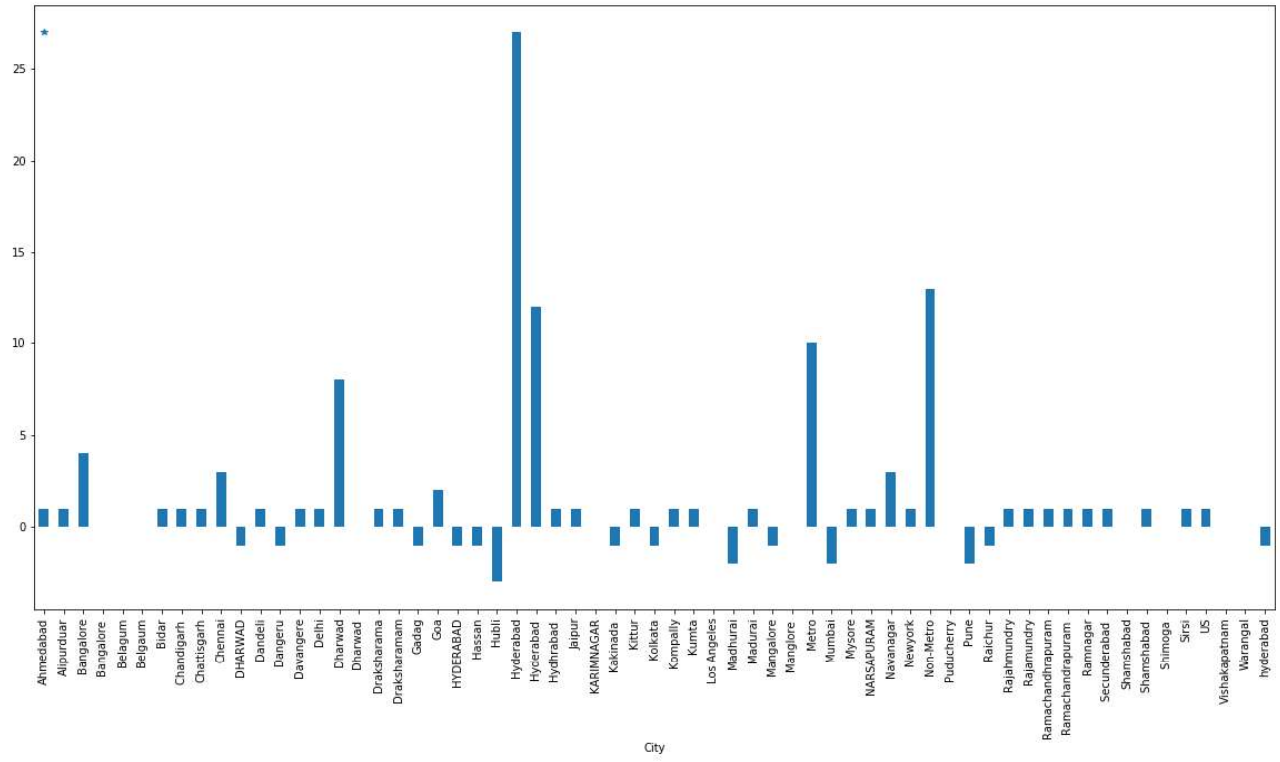
| City | |
|---|---|
| Ahmedabad | 1 |
| Alipurduar | 1 |
| Bangalore | 4 |
| Bangalore | 0 |
| Belagum | 0 |
| Belgaum | 0 |
| Bidar | 1 |
| Chandigarh | 1 |
| Chattisgarh | 1 |
| Chennai | 3 |
| DHARWAD | -1 |
| Dandeli | 1 |
| Dangeru | -1 |
| Davangere | 1 |
| Delhi | 1 |
| Dharwad | 8 |
| Dharwad | 0 |
| Draksharama | 1 |
| Draksharamam | 1 |
| Gadag | -1 |
| Goa | 2 |
| HYDERABAD | -1 |
| Hassan | -1 |
| Hubli | -3 |
| Hyderabad | 27 |
| Hyderabad | 12 |
| Hydhrabad | 1 |
| Jaipur | 1 |
| KARIMNAGAR | 0 |
| Kakinada | -1 |
| | .. |
| Kumta | 1 |
| Los Angeles | 0 |
| Madhurai | -2 |
| Madurai | 1 |
| Mangalore | -1 |
| Manglore | 0 |
| Metro | 10 |
| Mumbai | -2 |
| Mysore | 1 |

```
NARSAPURAM            1
Navanagar             3
Newyork               1
Non-Metro            13
Puducherry            0
Pune                 -2
Raichur              -1
Rajahmundry           1
Rajamundry            1
Ramachandhrapuram     1
Ramachandrapuram      1
Ramnagar              1
Secunderabad          1
Shamshabad            0
Shamshabad            1
Shimoga               0
Sirsi                 1
US                    1
Vishakapatnam         0
Warangal              0
hyderabad            -1
Name: City_preference, Length: 63, dtype: int64
```

In [14]:
```
print("The city with Highest Entrepreneurship is")
print(k)
fig = plt.figure(figsize=(20, 10))
City_more_preference.plot(kind='bar')
plt.plot(k,marker='*')
```

```
The city with Highest Entrepreneurship is
          City_more_preference
City
Hyderabad                    27
```

Out[14]: [<matplotlib.lines.Line2D at 0x29698cb2278>]



In [15]:

```
#college tier to role mapping
df["role"]=df['Work area']
dk=pd.DataFrame()
i=0
while i<n:
    if(df["role"][i]!="BPO" and df["role"][i]!="Mechanical" and df["role"][i]!="IT & Software" and df["role"][i]!="HR"
):
        df["role"][i]="other"
    i=i+1
df["role"]=df["role"].replace("BPO",1)
df["role"]=df["role"].replace("HR",1000)
df["role"]=df["role"].replace("IT & Software",1000000)
df["role"]=df["role"].replace("Mechanical",1000000000)
df["role"]=df["role"].replace("other",1000000000000)
asd=df.groupby("Work area").Serial.sum()
df["College"].fillna('other', inplace = True)
dk["College_JobRole"]=df.groupby("College").role.sum()
i=0
dk["IT & Software"]=0
dk["BPO"]=0
dk["Mechanical"]=0
dk["HR"]=0
dk["other"]=0
dk["Serial"]=1
nk=dk["Serial"].sum()
del dk["Serial"]
while i<nk:
    a=dk["College_JobRole"][i]
    b=a%1000
    a=a/1000
    a=int(a)
    dk["BPO"][i]=b
    b=a%1000
    a=int(a/1000)
    dk["HR"][i]=b
    b=a%1000
    a=int(a/1000)
    dk["IT & Software"][i]=b
    b=a%1000
    a=int(a/1000)
    dk["Mechanical"][i]=b
```

```
    b=a%1000
    a=int(a/1000)
    dk["other"][i]=b
    i=i+1
del dk["College_JobRole"]
dk
```

Out[15]:

| College | IT & Software | BPO | Mechanical | HR | other |
|---|---|---|---|---|---|
| Alur venkta roa | 2 | 0 | 0 | 0 | 0 |
| BITS PILANI HYB CAMPUS | 1 | 0 | 0 | 0 | 0 |
| BVB | 2 | 3 | 0 | 7 | 8 |
| Basel Mission | 1 | 2 | 0 | 7 | 0 |
| Basel mission | 0 | 2 | 0 | 1 | 0 |
| Bms | 0 | 0 | 0 | 3 | 3 |
| Bms | 1 | 0 | 0 | 0 | 0 |
| CVR College of engineering | 1 | 0 | 0 | 0 | 0 |
| Ellenki institute of engineering and technology | 1 | 0 | 0 | 0 | 0 |
| GIT | 1 | 3 | 0 | 2 | 5 |
| Geethanjali College of Engineering and Technology | 1 | 0 | 0 | 0 | 0 |
| Googte | 1 | 0 | 0 | 0 | 0 |
| Govt College | 0 | 0 | 0 | 2 | 0 |
| Haliyal | 3 | 2 | 0 | 0 | 0 |
| JSS | 2 | 6 | 0 | 7 | 8 |
| Jnnc | 1 | 0 | 0 | 0 | 0 |
| KCD | 3 | 0 | 0 | 3 | 0 |
| KITE | 1 | 0 | 0 | 0 | 0 |
| KLE | 0 | 2 | 0 | 3 | 4 |
| KLEIT | 2 | 5 | 0 | 2 | 0 |
| Karavali institute of technology | 1 | 0 | 0 | 0 | 0 |
| Kcd | 1 | 0 | 0 | 1 | 0 |
| Keshav Memorial Institute of Technology | 0 | 0 | 0 | 0 | 1 |
| MGIT | 1 | 0 | 0 | 0 | 0 |

| College | IT & Software | BPO | Mechanical | HR | other |
|---|---|---|---|---|---|
| MVSR ENGINEERING COLLEGE | 0 | 0 | 0 | 0 | 1 |
| MVSR engineering college | 1 | 0 | 0 | 0 | 0 |
| Mallareddy Engineering college for women's | 1 | 0 | 0 | 0 | 0 |
| Mallareddy engineering college for women | 1 | 0 | 0 | 0 | 0 |
| New Pragathi college of commerce and science | 0 | 0 | 0 | 0 | 1 |
| Nmam | 1 | 4 | 0 | 2 | 2 |
| ... | ... | ... | ... | ... | ... |
| V.S.M. college of engineering,Ramachandrapuram. | 0 | 0 | 0 | 0 | 1 |
| VARDHAMAN COLLEGE OF ENGINEERING | 1 | 0 | 0 | 0 | 0 |
| VDRIT | 1 | 2 | 0 | 3 | 7 |
| VIGNANA BHARATHI INSTITUTE of TECHNOLOGY | 1 | 0 | 0 | 0 | 0 |
| VNR VJIET | 1 | 0 | 0 | 0 | 0 |
| VSM college of engineering,Ramchandrapuram | 0 | 0 | 0 | 1 | 0 |
| Vaagdevi college of engineering | 0 | 0 | 0 | 0 | 1 |
| Vardhaman | 0 | 0 | 0 | 0 | 1 |
| Vardhaman College | 1 | 0 | 0 | 0 | 0 |
| Vardhaman College Of Engineering | 0 | 0 | 0 | 0 | 1 |
| Vardhaman College of Engineering | 2 | 0 | 0 | 0 | 2 |
| Vardhaman College of Engineering | 2 | 0 | 0 | 0 | 1 |
| Vardhaman College of Engineering. | 1 | 0 | 0 | 0 | 0 |
| Vardhaman College of engineering | 1 | 0 | 0 | 0 | 0 |
| Vardhaman college | 0 | 0 | 0 | 1 | 0 |
| Vardhaman college of Engineering | 1 | 0 | 0 | 0 | 0 |
| Vardhaman college of engineering | 18 | 0 | 0 | 1 | 8 |
| Vardhaman college of engineering | 20 | 0 | 0 | 1 | 8 |

| College | IT & Software | BPO | Mechanical | HR | other |
|---|---|---|---|---|---|
| Vardhman college of engineering | 1 | 0 | 0 | 0 | 0 |
| Vbit | 0 | 0 | 0 | 0 | 2 |
| Vidyasagar University | 1 | 0 | 0 | 0 | 0 |
| Vignana Bharathi Institute of Technology | 1 | 0 | 0 | 0 | 0 |
| Vignana Bharathi institute of technology | 2 | 0 | 0 | 0 | 0 |
| Vsm college of engineering | 2 | 0 | 0 | 0 | 0 |
| Vsm college of engineering | 4 | 0 | 0 | 0 | 0 |
| Y. N. Degree college | 0 | 0 | 0 | 0 | 1 |
| institute of aeronautical engineering college | 1 | 0 | 0 | 0 | 0 |
| kle | 0 | 1 | 0 | 0 | 0 |
| other | 2 | 1 | 0 | 1 | 39 |
| siddhartha institute of science and technology | 1 | 0 | 0 | 0 | 0 |

78 rows × 5 columns

In [16]:

```
#7. Job role to stream
dp=pd.DataFrame()
dp["Stream_College"]=df.groupby("Stream").role.sum()
i=0
dp["IT & Software"]=0
dp["BPO"]=0
dp["Mechanical"]=0
dp["HR"]=0
dp["other"]=0
dp["Serial"]=1
nk=dp["Serial"].sum()
print(nk)
del dp["Serial"]
while i<nk:
    a=dp["Stream_College"][i]
    b=a%1000
    a=a/1000
    a=int(a)
    dp["BPO"][i]=b
    b=a%1000
    a=int(a/1000)
    dp["HR"][i]=b
    b=a%1000
    a=int(a/1000)
    dp["IT & Software"][i]=b
    b=a%1000
    a=int(a/1000)
    dp["Mechanical"][i]=b
    b=a%1000
    a=int(a/1000)
    dp["other"][i]=b
    i=i+1
del dp["Stream_College"]
dp
```

5

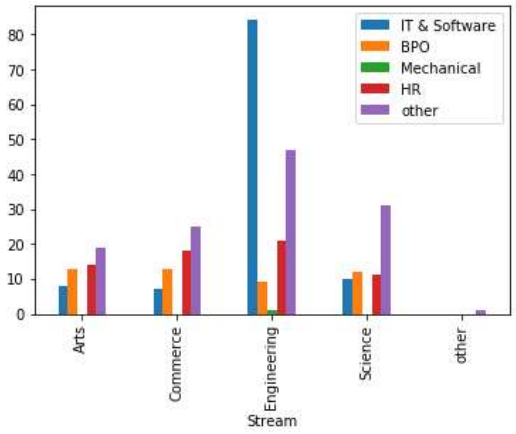
Out[16]:

| Stream | IT & Software | BPO | Mechanical | HR | other |
|---|---|---|---|---|---|
| **Arts** | 8 | 13 | 0 | 14 | 19 |
| **Commerce** | 7 | 13 | 0 | 18 | 25 |
| **Engineering** | 84 | 9 | 1 | 21 | 47 |
| **Science** | 10 | 12 | 0 | 11 | 31 |
| **other** | 0 | 0 | 0 | 0 | 1 |

In [17]:
```
fig = plt.figure(figsize=(15, 10))
dp.plot(kind='bar')
```

Out[17]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x296997c2c88>

<Figure size 1080x720 with 0 Axes>
```



In [ ]:

In [ ]:

# Methodology to set up and conduct a survey

## What steps will you follow to set up and conduct a survey?

The steps to be followed for survey are
i) Building a team of different regions
ii) Selection of Question
iii) Selection of survey method (Online and offline)
IV) Selection of media for online survey
v) Analyzing the survey
vi) Analyzing the survey according to region
vii) Plotting the Graph
viii) Write on survey
ix) Submitting report

**Building a team:**

A team from different regions should be selected so that we get information from all regions.

**Selection of Question:**

We need to select proper question which will help for analyzing, it should be MCQ question. It will be easy to analyze the data

**Selection of survey method:**

We need to select survey which will be low cost and give proper answer. Offline survey should be taken by candidates by asking suitable person

### How did you set up the questionnaire? What criteria you used to frame questions?

We need to select a question which will help for our survey
The criteria to be seen are:
i)Question should be relevant to survey
ii) No Grammar mistake in Question
iii) It must be MCQ
iv) Easy to understand
v) Less question but more information

## Any issue or obstacle you have faced in the survey process

Yeah, they are many obstacles to be faced in the survey
i) Google survey is too costly to afford (1$ per person, if we need 500 survey than 68rs*500)
ii) No person likes to give answer for free we need to offer them something for them
iii) People will lie for few questions
iv) Can’t collect huge data offline.