

Adv Python - Module 2 Practice - Anna Christensen -

March 25, 2025

```
[3]: import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

```
[4]: df = pd.read_csv('companies.csv')
df.head()
```

```
[4]:  company_name      url      city state      country  employees \
0      OpenAI      openai.com  San Francisco  CA  United States      655
1      Alchemy      alchemy.com  San Francisco  CA  United States      201
2      dbt Labs      getdbt.com  Philadelphia  PA  United States      511
3      Kindbody      kindbody.com      New York  NY  United States      621
4      Anthropic  anthropic.com  San Francisco  CA      USA      131

      founded      Industry  estimated_revenues  job_openings \
0      2015.0      AI      88722200.0      25.0
1      NaN      Fintech      39269750.0      1.0
2      2016.0      Analytics      76987400.0      1.0
3      2018.0  Digital Health      105791000.0      38.0
4      NaN      AI      15400000.0      1.0

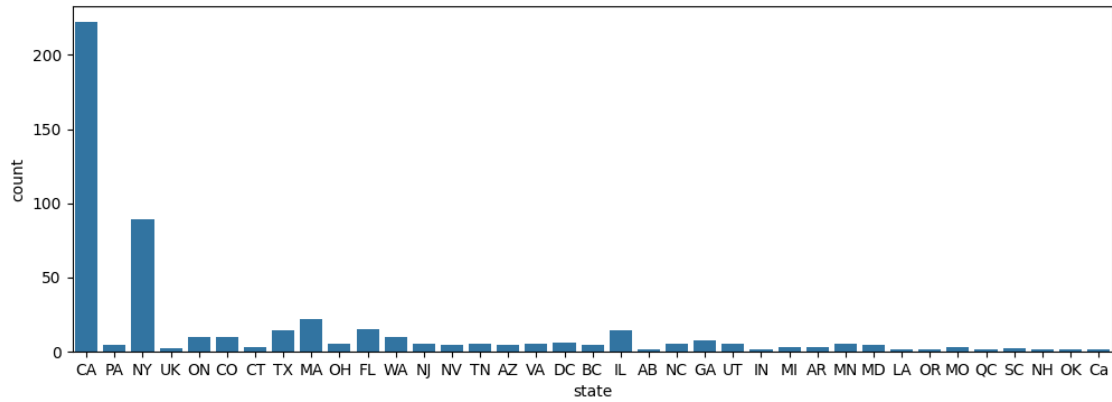
      keywords      LeadInvestors \
0  Artificial Intelligence, Online Gaming, Non Pr...  Thrive, Founders Fund
1      NaN  Lightspeed, Silver Lake
2      NaN      Altimeter
3  Healthcare, Medical, Women-Focused, Corporate ...      NaN
4      NaN      NaN

      Accelerator  btype      valuation  total_funding  growth_percentage
0      NaN      NaN  2.900000e+10  11000000000.0      1.18
1      NaN      NaN  1.020000e+10  564000000.0      1.34
2      NaN      NaN  4.200000e+09  414400000.0      0.76
3      NaN      NaN  1.800000e+09  125300000.0      0.44
4      NaN      NaN  4.400000e+09  1300000000.0      1.38
```

```
[18]: df['company_name'].value_counts()
df['state'].value_counts()
```

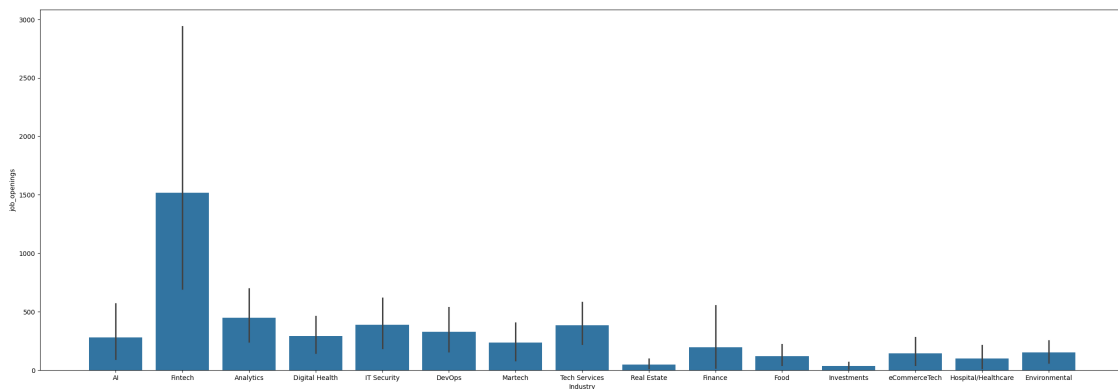
```
[18]: state
      CA      222
      NY      89
      MA      22
      FL      15
      TX      14
      IL      14
      ON      10
      CO      10
      WA      10
      GA       7
      DC       6
      TN       5
      VA       5
      OH       5
      UT       5
      NJ       5
      MN       5
      NC       5
      MD       4
      BC       4
      AZ       4
      NV       4
      PA       4
      MI       3
      AR       3
      CT       3
      MO       3
      UK       2
      SC       2
      QC       1
      OK       1
      NH       1
      AB       1
      OR       1
      LA       1
      IN       1
      Ca       1
      Name: count, dtype: int64
```

```
[27]: plt.figure(figsize=(12, 4))
      sns.countplot(x='state', data=df,)
      plt.show()
```



```
[43]: plt.figure(figsize=(30, 10))
sns.barplot(x='Industry', y='job_openings', data=df, estimator=sum)
plt.plot()
```

[43]: []



```
[46]: df.head(2)
```

```
[46]:   company_name      url      city state      country  employees \
0    OpenAI    openai.com  San Francisco    CA  United States    655
1    Alchemy    alchemy.com  San Francisco    CA  United States    201

   founded Industry  estimated_revenues  job_openings \
0   2015.0      AI      88722200.0      25.0
1     NaN  Fintech      39269750.0      1.0

                                   keywords      LeadInvestors \
0  Artificial Intelligence, Online Gaming, Non Pr...  Thrive, Founders Fund
```

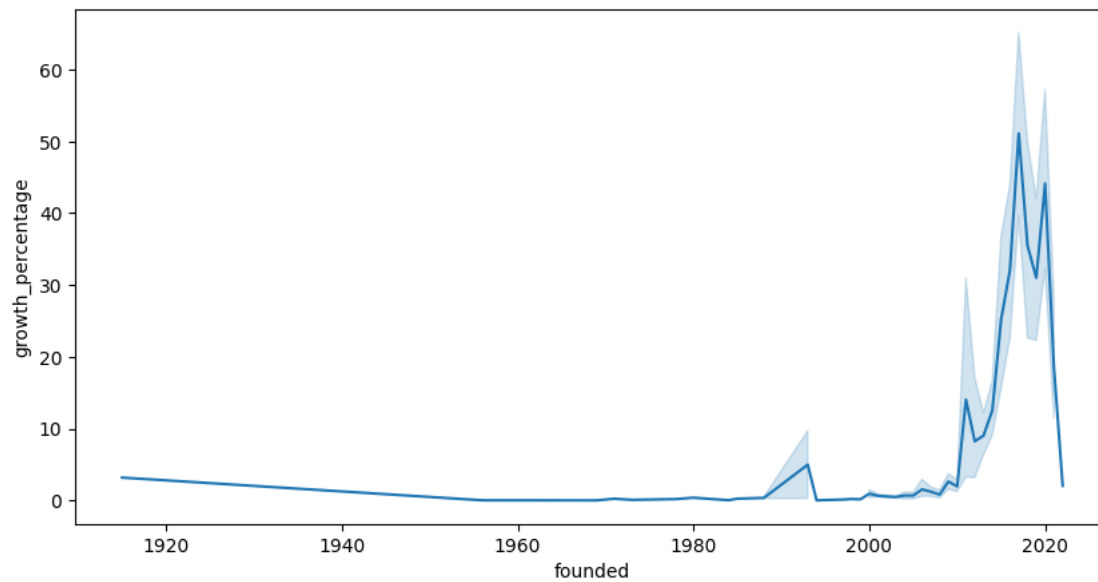
1

NaN Lightspeed, Silver Lake

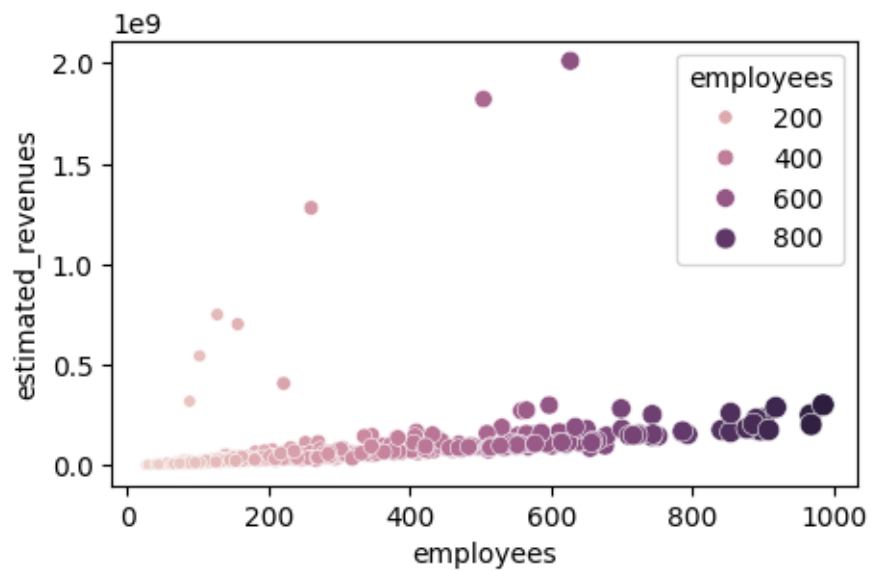
	Accelerator	btype	valuation	total_funding	growth_percentage
0	NaN	NaN	2.900000e+10	11000000000.0	1.18
1	NaN	NaN	1.020000e+10	564000000.0	1.34

```
[100]: plt.figure(figsize=(10, 5))
sns.lineplot(x='founded', y='growth_percentage', data=df, estimator=sum)
plt.plot()
```

[100]: []



```
[119]: plt.figure(figsize=(5, 3))
sns.scatterplot(x='employees', y='estimated_revenues', data=df,
               size='employees', hue='employees' )
plt.show()
```



```
[128]: plt.figure(figsize=(8, 3))
sns.histplot(x='founded', data=df)
plt.show()
```

