salary_exploration

February 25, 2023

[1]:
```
#read pickle file
import pickle
import pandas as pd
with open("../data/salary_guide.pkl", 'rb') as f:
     data = pickle.load(f)

#load to dataframe
df = pd.DataFrame(data)
```

[2]:
```
df
```

[2]:

	province	region	job	entry	mid	senior
0	alberta	calgary metropolitan region	administrative assistant	43.6-61.3	49.9-61.3	55.2-66.7
1	alberta	calgary metropolitan region	administrative manager	65.2-81.9	71.1-81.9	78.7-94.6
2	alberta	calgary metropolitan region	buyer assistant	45.5-66.1	53.4-66.1	58.8-72.2
3	alberta	calgary metropolitan region	claims supervisor	70.1-104.9	78.5-104.9	89.0-119.7
4	alberta	calgary metropolitan region	contracts administrator	80.4-109.3	87.7-109.3	91.2-120.5
…	…	…	…	…	…	…
18600	saskatchewan	saskatoon metropolitan area	project coordinator	65.3-92.4	72.7-92.4	81.9-104.2
18601	saskatchewan	saskatoon metropolitan area	project manager (infrastructure)	95.9-150.5	108.8-150.5	124.8-172.7
18602	saskatchewan	saskatoon metropolitan area	project manager (software)	95.0-144.0	106.4-144.0	120.5-163.4
18603	saskatchewan	saskatoon metropolitan area	scrum master	72.7-109.5	82.3-109.5	93.9-125.0
18604	saskatchewan	saskatoon metropolitan area	technical writer	58.9-91.3	68.9-91.3	80.3-103.9

[18605 rows x 6 columns]

[3]:
```
# Split the salary ranges into minimum and maximum salaries
df[['entry_min', 'entry_max']] = df['entry'].str.split('-', expand=True).
 ↪astype(float)
df[['mid_min', 'mid_max']] = df['mid'].str.split('-', expand=True).astype(float)
df[['senior_min', 'senior_max']] = df['senior'].str.split('-', expand=True).
 ↪astype(float)

# Create a new column with the maximum salary for each job
df['max_salary'] = df[['entry_max', 'mid_max', 'senior_max']].max(axis=1)
#calculate the mean for each level
df['entry_mean'] = df[['entry_min', 'entry_max']].mean(axis=1)
df['mid_mean'] = df[['mid_min', 'mid_max']].mean(axis=1)
df['senior_mean'] = df[['senior_min', 'senior_max']].mean(axis=1)

# Sort the DataFrame by the maximum salary in descending order
df_sorted = df.sort_values('max_salary', ascending=False)

# Get the job that pays the most
job_highest_paid = df_sorted['job'].iloc[0]

print(f"The job that pays the most is {job_highest_paid}")
```

The job that pays the most is receptionist (bilingual)

[4]:
```
df_sorted.head()
```

[4]:

	province	region	job	entry	mid	senior	entry_min	entry_max	mid_min	mid_max	senior_min	senior_max
9603	ontario	simcoe county	receptionist (bilingual)	50.3-66.8	56.1-66.8	63.1-713.0	50.3	66.8	56.1	66.8	63.1	713.0
6861	ontario	greater toronto area	chief financial officer	172.6-430.1	250.7-430.1	346.3-594.1	172.6	430.1	250.7	430.1	346.3	594.1
11160	ontario	york region	chief financial officer	170.0-424.8	246.7-424.8	341.5-587.9	170.0	424.8	246.7	424.8	341.5	587.9
215	alberta	calgary metropolitan region	chief financial officer	161.5-414.0	236.8-414.0	330.7-579.1	161.5	414.0	236.8	414.0	330.7	579.1
1967	british columbia	greater vancouver	chief financial officer	163.7-414.8	238.5-414.8	332.1-577.7						

```
1967       163.7      414.8    238.5    414.8       332.1       577.7

       max_salary  entry_mean  mid_mean  senior_mean
9603        713.0       58.55     61.45       388.05
6861        594.1      301.35    340.40       470.20
11160       587.9      297.40    335.75       464.70
215         579.1      287.75    325.40       454.90
1967        577.7      289.25    326.65       454.90
```

[5]:
```
df_sorted.head(100)
```

[5]:
```
               province                        region  \
9603            ontario                 simcoe county
6861            ontario          greater toronto area
11160           ontario                   york region
215             alberta  calgary metropolitan region
1967   british columbia             greater vancouver
…                 …                             …
11333           ontario                   york region
8514            ontario    ottawa metropolitan region
8534            ontario    ottawa metropolitan region
5514            ontario                 durham region
5534            ontario                 durham region

                           job        entry          mid       senior  \
9603   receptionist (bilingual)    50.3-66.8    56.1-66.8   63.1-713.0
6861    chief financial officer  172.6-430.1  250.7-430.1  346.3-594.1
11160   chief financial officer  170.0-424.8  246.7-424.8  341.5-587.9
215     chief financial officer  161.5-414.0  236.8-414.0  330.7-579.1
1967    chief financial officer  163.7-414.8  238.5-414.8  332.1-577.7
…                          …          …           …           …
11333      sales vice-president  144.3-278.3  194.8-278.3  251.1-358.7
8514   marketing vice-president  129.4-277.3  180.8-277.3  232.7-357.0
8534       sales vice-president  144.0-276.8  193.8-276.8  249.8-356.9
5514   marketing vice-president  128.5-276.3  179.8-276.3  231.7-356.2
5534       sales vice-president  143.0-275.9  192.9-275.9  248.8-356.1

       entry_min  entry_max  mid_min  mid_max  senior_min  senior_max  \
9603        50.3       66.8     56.1     66.8        63.1       713.0
6861       172.6      430.1    250.7    430.1       346.3       594.1
11160      170.0      424.8    246.7    424.8       341.5       587.9
215        161.5      414.0    236.8    414.0       330.7       579.1
1967       163.7      414.8    238.5    414.8       332.1       577.7
…            …          …        …        …           …           …
11333      144.3      278.3    194.8    278.3       251.1       358.7
8514       129.4      277.3    180.8    277.3       232.7       357.0
8534       144.0      276.8    193.8    276.8       249.8       356.9
```

```
5514       128.5      276.3    179.8    276.3       231.7       356.2
5534       143.0      275.9    192.9    275.9       248.8       356.1

       max_salary  entry_mean  mid_mean  senior_mean
9603        713.0       58.55     61.45       388.05
6861        594.1      301.35    340.40       470.20
11160       587.9      297.40    335.75       464.70
215         579.1      287.75    325.40       454.90
1967        577.7      289.25    326.65       454.90
…             …           …         …            …
11333       358.7      211.30    236.55       304.90
8514        357.0      203.35    229.05       294.85
8534        356.9      210.40    235.30       303.35
5514        356.2      202.40    228.05       293.95
5534        356.1      209.45    234.40       302.45

[100 rows x 16 columns]
```

0.0.1 Get the engineering jobs that pay the most

[6]:
```
top_engineers_jobs=df_sorted[df_sorted['job'].str.contains('engineer',␣
↪case=False)]
top_engineers_jobs.head(30)
```

[6]:
```
              province                        region                 job  \
6784           ontario         greater toronto area  engineering manager
6968           ontario         greater toronto area  engineering manager
6710           ontario         greater toronto area  engineering manager
11123          ontario                  york region  engineering manager
11266          ontario                  york region  engineering manager
5295           ontario                durham region  engineering manager
5221           ontario                durham region  engineering manager
5466           ontario                durham region  engineering manager
938            alberta             northern alberta  engineering manager
1087           alberta             northern alberta  engineering manager
867            alberta             northern alberta  engineering manager
1486           alberta                fort st. john  engineering manager
8990           ontario                  peel region  engineering manager
9064           ontario                  peel region  engineering manager
9245           ontario                  peel region  engineering manager
136            alberta  calgary metropolitan region  engineering manager
65             alberta  calgary metropolitan region  engineering manager
318            alberta  calgary metropolitan region  engineering manager
12234           québec         chaudière-appalaches  engineering manager
12017           québec         chaudière-appalaches  engineering manager
12085           québec         chaudière-appalaches  engineering manager
16162           québec                     montréal  engineering manager
```

16231	québec	montréal	engineering manager	
16400	québec	montréal	engineering manager	
1072	alberta	northern alberta	project engineer	
988	alberta	northern alberta	project engineer	
889	alberta	northern alberta	project engineer	
3088	british columbia	surrey/delta	engineering manager	
3207	british columbia	surrey/delta	engineering manager	
3017	british columbia	surrey/delta	engineering manager	

	entry	mid	senior	entry_min	entry_max	mid_min	\
6784	101.4-157.2	115.8-157.2	133.0-178.0	101.4	157.2	115.8	
6968	101.4-157.2	115.8-157.2	133.0-178.0	101.4	157.2	115.8	
6710	101.4-157.2	115.8-157.2	133.0-178.0	101.4	157.2	115.8	
11123	99.3-154.1	113.5-154.1	130.4-174.4	99.3	154.1	113.5	
11266	99.3-154.1	113.5-154.1	130.4-174.4	99.3	154.1	113.5	
5295	98.9-153.2	112.9-153.2	129.7-173.1	98.9	153.2	112.9	
5221	98.9-153.2	112.9-153.2	129.7-173.1	98.9	153.2	112.9	
5466	98.9-153.2	112.9-153.2	129.7-173.1	98.9	153.2	112.9	
938	98.2-151.3	111.8-151.3	128.0-171.4	98.2	151.3	111.8	
1087	98.2-151.3	111.8-151.3	128.0-171.4	98.2	151.3	111.8	
867	98.2-151.3	111.8-151.3	128.0-171.4	98.2	151.3	111.8	
1486	88.2-137.1	100.5-137.1	115.5-170.0	88.2	137.1	100.5	
8990	94.5-147.1	108.0-147.1	124.3-166.5	94.5	147.1	108.0	
9064	94.5-147.1	108.0-147.1	124.3-166.5	94.5	147.1	108.0	
9245	94.5-147.1	108.0-147.1	124.3-166.5	94.5	147.1	108.0	
136	94.0-146.3	107.3-146.3	123.3-166.4	94.0	146.3	107.3	
65	94.0-146.3	107.3-146.3	123.3-166.4	94.0	146.3	107.3	
318	94.0-146.3	107.3-146.3	123.3-166.4	94.0	146.3	107.3	
12234	94.0-146.3	107.3-146.3	123.3-166.4	94.0	146.3	107.3	
12017	94.0-146.3	107.3-146.3	123.3-166.4	94.0	146.3	107.3	
12085	94.0-146.3	107.3-146.3	123.3-166.4	94.0	146.3	107.3	
16162	93.4-146.4	107.1-146.4	123.4-166.1	93.4	146.4	107.1	
16231	93.4-146.4	107.1-146.4	123.4-166.1	93.4	146.4	107.1	
16400	93.4-146.4	107.1-146.4	123.4-166.1	93.4	146.4	107.1	
1072	105.9-148.1	114.5-148.1	125.8-164.2	105.9	148.1	114.5	
988	105.9-148.1	114.5-148.1	125.8-164.2	105.9	148.1	114.5	
889	105.9-148.1	114.5-148.1	125.8-164.2	105.9	148.1	114.5	
3088	93.9-145.0	107.0-145.0	122.8-164.1	93.9	145.0	107.0	
3207	93.9-145.0	107.0-145.0	122.8-164.1	93.9	145.0	107.0	
3017	93.9-145.0	107.0-145.0	122.8-164.1	93.9	145.0	107.0	

	mid_max	senior_min	senior_max	max_salary	entry_mean	mid_mean	\
6784	157.2	133.0	178.0	178.0	129.30	136.50	
6968	157.2	133.0	178.0	178.0	129.30	136.50	
6710	157.2	133.0	178.0	178.0	129.30	136.50	
11123	154.1	130.4	174.4	174.4	126.70	133.80	
11266	154.1	130.4	174.4	174.4	126.70	133.80	

5295	153.2	129.7	173.1	173.1	126.05	133.05
5221	153.2	129.7	173.1	173.1	126.05	133.05
5466	153.2	129.7	173.1	173.1	126.05	133.05
938	151.3	128.0	171.4	171.4	124.75	131.55
1087	151.3	128.0	171.4	171.4	124.75	131.55
867	151.3	128.0	171.4	171.4	124.75	131.55
1486	137.1	115.5	170.0	170.0	112.65	118.80
8990	147.1	124.3	166.5	166.5	120.80	127.55
9064	147.1	124.3	166.5	166.5	120.80	127.55
9245	147.1	124.3	166.5	166.5	120.80	127.55
136	146.3	123.3	166.4	166.4	120.15	126.80
65	146.3	123.3	166.4	166.4	120.15	126.80
318	146.3	123.3	166.4	166.4	120.15	126.80
12234	146.3	123.3	166.4	166.4	120.15	126.80
12017	146.3	123.3	166.4	166.4	120.15	126.80
12085	146.3	123.3	166.4	166.4	120.15	126.80
16162	146.4	123.4	166.1	166.1	119.90	126.75
16231	146.4	123.4	166.1	166.1	119.90	126.75
16400	146.4	123.4	166.1	166.1	119.90	126.75
1072	148.1	125.8	164.2	164.2	127.00	131.30
988	148.1	125.8	164.2	164.2	127.00	131.30
889	148.1	125.8	164.2	164.2	127.00	131.30
3088	145.0	122.8	164.1	164.1	119.45	126.00
3207	145.0	122.8	164.1	164.1	119.45	126.00
3017	145.0	122.8	164.1	164.1	119.45	126.00

	senior_mean
6784	155.50
6968	155.50
6710	155.50
11123	152.40
11266	152.40
5295	151.40
5221	151.40
5466	151.40
938	149.70
1087	149.70
867	149.70
1486	142.75
8990	145.40
9064	145.40
9245	145.40
136	144.85
65	144.85
318	144.85
12234	144.85
12017	144.85

```
12085       144.85
16162       144.75
16231       144.75
16400       144.75
1072        145.00
988         145.00
889         145.00
3088        143.45
3207        143.45
3017        143.45
```

0.0.2 get the engineers jobs that pay the least

[7]:
```
worst_engineers_jobs=top_engineers_jobs.sort_values('max_salary',␣
↪ascending=True)
worst_engineers_jobs.head(30)
```

[7]:
```
      province                          region  \
10757  ontario           windsor-essex county
6305   ontario           greater sudbury area
4754   ontario                   brant county
5873   ontario          greater hamilton area
6744   ontario           greater toronto area
12398   québec                         estrie
12337   québec                         estrie
11659   québec                centre-du-québec
14708   québec  montérégie - brome-missisquoi
11598   québec                centre-du-québec
14769   québec  montérégie - brome-missisquoi
12737   québec                      lanaudière
12584   québec                         estrie
12474   québec                         estrie
15521   québec   montérégie - les maskoutains
15460   québec   montérégie - les maskoutains
14844   québec  montérégie - brome-missisquoi
12397   québec                         estrie
14919   québec  montérégie - brome-missisquoi
12336   québec                         estrie
11735   québec                centre-du-québec
11850   québec                centre-du-québec
15126   québec  montérégie - la haute-yamaska
15064   québec  montérégie - la haute-yamaska
12814   québec                      lanaudière
11597   québec                centre-du-québec
14707   québec  montérégie - brome-missisquoi
14768   québec  montérégie - brome-missisquoi
11658   québec                centre-du-québec
```

```
12736   québec                        lanaudière

                                        job      entry        mid     senior  \
10757  structural engineering technologist  55.0-65.0  60.0-65.0  65.0-75.0
6305   structural engineering technologist  55.0-65.0  60.0-65.0  65.0-75.0
4754   structural engineering technologist  55.0-65.0  60.0-65.0  65.0-75.0
5873   structural engineering technologist  55.0-65.0  60.0-65.0  65.0-75.0
6744   structural engineering technologist  55.0-65.0  60.0-65.0  65.0-75.0
12398       civil engineering technologist  41.4-67.2  49.2-67.2  57.6-77.0
12337       civil engineering technologist  41.4-67.2  49.2-67.2  57.6-77.0
11659       civil engineering technologist  41.3-67.5  49.2-67.5  57.8-77.4
14708       civil engineering technologist  41.7-67.6  49.5-67.6  58.0-77.4
11598       civil engineering technologist  41.3-67.5  49.2-67.5  57.8-77.4
14769       civil engineering technologist  41.7-67.6  49.5-67.6  58.0-77.4
12737       civil engineering technologist  41.8-67.7  49.6-67.7  58.1-77.5
12584                  stationary engineer  48.4-69.7  53.3-69.7  59.3-77.8
12474                  stationary engineer  48.4-69.7  53.3-69.7  59.3-77.8
15521       civil engineering technologist  42.2-68.4  50.1-68.4  58.7-78.3
15460       civil engineering technologist  42.2-68.4  50.1-68.4  58.7-78.3
14844                  stationary engineer  48.6-70.3  53.6-70.3  59.7-78.4
12397            civil engineering drafter  47.8-68.8  53.8-68.8  61.0-78.4
14919                  stationary engineer  48.6-70.3  53.6-70.3  59.7-78.4
12336            civil engineering drafter  47.8-68.8  53.8-68.8  61.0-78.4
11735                  stationary engineer  48.5-70.3  53.5-70.3  59.7-78.5
11850                  stationary engineer  48.5-70.3  53.5-70.3  59.7-78.5
15126       civil engineering technologist  42.2-68.6  50.2-68.6  58.8-78.5
15064       civil engineering technologist  42.2-68.6  50.2-68.6  58.8-78.5
12814                  stationary engineer  48.8-70.6  53.8-70.6  59.9-78.7
11597            civil engineering drafter  47.8-69.1  53.8-69.1  61.1-78.8
14707            civil engineering drafter  48.2-69.2  54.1-69.2  61.4-78.8
14768            civil engineering drafter  48.2-69.2  54.1-69.2  61.4-78.8
11658            civil engineering drafter  47.8-69.1  53.8-69.1  61.1-78.8
12736            civil engineering drafter  48.3-69.3  54.2-69.3  61.4-78.9

       entry_min  entry_max  mid_min  mid_max  senior_min  senior_max  \
10757       55.0       65.0     60.0     65.0        65.0        75.0
6305        55.0       65.0     60.0     65.0        65.0        75.0
4754        55.0       65.0     60.0     65.0        65.0        75.0
5873        55.0       65.0     60.0     65.0        65.0        75.0
6744        55.0       65.0     60.0     65.0        65.0        75.0
12398       41.4       67.2     49.2     67.2        57.6        77.0
12337       41.4       67.2     49.2     67.2        57.6        77.0
11659       41.3       67.5     49.2     67.5        57.8        77.4
14708       41.7       67.6     49.5     67.6        58.0        77.4
11598       41.3       67.5     49.2     67.5        57.8        77.4
14769       41.7       67.6     49.5     67.6        58.0        77.4
12737       41.8       67.7     49.6     67.7        58.1        77.5
```

12584	48.4	69.7	53.3	69.7	59.3	77.8
12474	48.4	69.7	53.3	69.7	59.3	77.8
15521	42.2	68.4	50.1	68.4	58.7	78.3
15460	42.2	68.4	50.1	68.4	58.7	78.3
14844	48.6	70.3	53.6	70.3	59.7	78.4
12397	47.8	68.8	53.8	68.8	61.0	78.4
14919	48.6	70.3	53.6	70.3	59.7	78.4
12336	47.8	68.8	53.8	68.8	61.0	78.4
11735	48.5	70.3	53.5	70.3	59.7	78.5
11850	48.5	70.3	53.5	70.3	59.7	78.5
15126	42.2	68.6	50.2	68.6	58.8	78.5
15064	42.2	68.6	50.2	68.6	58.8	78.5
12814	48.8	70.6	53.8	70.6	59.9	78.7
11597	47.8	69.1	53.8	69.1	61.1	78.8
14707	48.2	69.2	54.1	69.2	61.4	78.8
14768	48.2	69.2	54.1	69.2	61.4	78.8
11658	47.8	69.1	53.8	69.1	61.1	78.8
12736	48.3	69.3	54.2	69.3	61.4	78.9

	max_salary	entry_mean	mid_mean	senior_mean
10757	75.0	60.00	62.50	70.00
6305	75.0	60.00	62.50	70.00
4754	75.0	60.00	62.50	70.00
5873	75.0	60.00	62.50	70.00
6744	75.0	60.00	62.50	70.00
12398	77.0	54.30	58.20	67.30
12337	77.0	54.30	58.20	67.30
11659	77.4	54.40	58.35	67.60
14708	77.4	54.65	58.55	67.70
11598	77.4	54.40	58.35	67.60
14769	77.4	54.65	58.55	67.70
12737	77.5	54.75	58.65	67.80
12584	77.8	59.05	61.50	68.55
12474	77.8	59.05	61.50	68.55
15521	78.3	55.30	59.25	68.50
15460	78.3	55.30	59.25	68.50
14844	78.4	59.45	61.95	69.05
12397	78.4	58.30	61.30	69.70
14919	78.4	59.45	61.95	69.05
12336	78.4	58.30	61.30	69.70
11735	78.5	59.40	61.90	69.10
11850	78.5	59.40	61.90	69.10
15126	78.5	55.40	59.40	68.65
15064	78.5	55.40	59.40	68.65
12814	78.7	59.70	62.20	69.30
11597	78.8	58.45	61.45	69.95
14707	78.8	58.70	61.65	70.10

```
14768        78.8       58.70     61.65        70.10
11658        78.8       58.45     61.45        69.95
12736        78.9       58.80     61.75        70.15
```

0.0.3 Get the job that pays the most in Quebec

[8]:
```
top_quebec_jobs=df_sorted[df_sorted['province'].str.contains('québec',␣
↪case=False)]
top_quebec_jobs.head(30)
```

[8]:
```
      province                            region  \
16292   québec                          montréal
14480   québec       montérégie - agglomeration
13654   québec                             laval
17164   québec                         outaouais
17512   québec            saguenay-lac-saint-jean
14088   québec                          mauricie
12823   québec                        lanaudière
11744   québec                  centre-du-québec
13217   québec                       laurentides
16737   québec                  national capital
12483   québec                            estrie
12486   québec                            estrie
11748   québec                  centre-du-québec
11747   québec                  centre-du-québec
14092   québec                          mauricie
16401   québec                          montréal
13758   québec                             laval
16295   québec                          montréal
14483   québec       montérégie - agglomeration
13657   québec                             laval
16469   québec                          montréal
13827   québec                             laval
14655   québec       montérégie - agglomeration
14191   québec                          mauricie
15314   québec     montérégie - la haute-yamaska
16061   québec  montérégie - vaudreuil-soulanges
12588   québec                            estrie
15708   québec      montérégie - les maskoutains
14923   québec     montérégie - brome-missisquoi
17342   québec                         outaouais

                                job        entry          mid  \
16292          chief financial officer  161.6-413.3  236.8-413.3
14480          chief financial officer  160.5-410.4  235.3-410.4
13654          chief financial officer  159.4-409.3  234.0-409.3
17164          chief financial officer  141.3-380.8  211.4-380.8
```

```
17512           chief financial officer  138.4-377.3  208.3-377.3
14088           chief financial officer  142.7-375.2  207.1-375.2
12823           chief financial officer  146.8-226.3  162.2-226.3
11744           chief financial officer  140.0-370.1  203.5-370.1
13217           chief financial officer  135.6-369.9  203.6-369.9
16737           chief financial officer  128.3-367.5  198.9-367.5
12483           chief financial officer  140.0-369.3  203.5-369.3
12486            vice president finance  136.7-344.5  199.2-344.5
11748                        controller  136.7-344.5  199.2-344.5
11747            vice president finance  136.7-344.5  199.2-344.5
14092            vice president finance  136.7-344.5  199.2-344.5
16401     head of workshop and assembly  168.1-328.4  237.2-328.4
13758     head of workshop and assembly  165.9-325.1  234.4-325.1
16295            vice president finance  150.6-289.7  167.1-289.7
14483            vice president finance  149.7-289.6  166.9-289.6
13657            vice president finance  143.8-286.6  164.8-286.6
16469  vice-president marketing & sales  119.6-294.8  164.5-294.8
13827  vice-president marketing & sales  117.8-291.7  162.3-291.7
14655  vice-president marketing & sales  118.0-291.4  162.4-291.4
14191     head of workshop and assembly  144.2-294.8  208.3-294.8
15314     head of workshop and assembly  143.1-293.1  207.0-293.1
16061  vice-president marketing & sales  117.5-289.2  161.1-289.2
12588     head of workshop and assembly  142.0-292.4  206.0-292.4
15708     head of workshop and assembly  142.9-292.0  206.2-292.0
14923     head of workshop and assembly  141.9-291.5  205.6-291.5
17342  vice-president marketing & sales  115.2-286.0  158.6-286.0

            senior  entry_min  entry_max  mid_min  mid_max  senior_min  \
16292  330.5-576.4      161.6      413.3    236.8    413.3       330.5
14480  328.3-571.8      160.5      410.4    235.3    410.4       328.3
13654  327.1-571.4      159.4      409.3    234.0    409.3       327.1
17164  300.6-541.1      141.3      380.8    211.4    380.8       300.6
17512  297.2-537.6      138.4      377.3    208.3    377.3       297.2
14088  295.6-534.6      142.7      375.2    207.1    375.2       295.6
12823  293.4-531.8      146.8      226.3    162.2    226.3       293.4
11744  293.9-528.3      140.0      370.1    203.5    370.1       293.9
13217  291.1-528.2      135.6      369.9    203.6    369.9       291.1
16737  287.7-527.7      128.3      367.5    198.9    367.5       287.7
12483  293.9-527.6      140.0      369.3    203.5    369.3       293.9
12486  269.2-506.0      136.7      344.5    199.2    344.5       269.2
11748  269.2-506.0      136.7      344.5    199.2    344.5       269.2
11747  269.2-506.0      136.7      344.5    199.2    344.5       269.2
14092  269.2-506.0      136.7      344.5    199.2    344.5       269.2
16401  297.2-411.3      168.1      328.4    237.2    328.4       297.2
13758  294.0-407.4      165.9      325.1    234.4    325.1       294.0
16295  231.4-401.4      150.6      289.7    167.1    289.7       231.4
14483  230.4-400.5      149.7      289.6    166.9    289.6       230.4
```

```
13657  228.7-397.6      143.8      286.6    164.8    286.6       228.7
16469  211.7-379.5      119.6      294.8    164.5    294.8       211.7
13827  209.2-375.7      117.8      291.7    162.3    291.7       209.2
14655  209.2-375.1      118.0      291.4    162.4    291.4       209.2
14191  265.0-374.7      144.2      294.8    208.3    294.8       265.0
15314  263.4-372.7      143.1      293.1    207.0    293.1       263.4
16061  207.5-372.5      117.5      289.2    161.1    289.2       207.5
12588  262.6-372.3      142.0      292.4    206.0    292.4       262.6
15708  262.4-371.3      142.9      292.0    206.2    292.0       262.4
14923  261.8-370.9      141.9      291.5    205.6    291.5       261.8
17342  204.7-369.2      115.2      286.0    158.6    286.0       204.7

       senior_max  max_salary  entry_mean  mid_mean  senior_mean
16292       576.4       576.4      287.45    325.05       453.45
14480       571.8       571.8      285.45    322.85       450.05
13654       571.4       571.4      284.35    321.65       449.25
17164       541.1       541.1      261.05    296.10       420.85
17512       537.6       537.6      257.85    292.80       417.40
14088       534.6       534.6      258.95    291.15       415.10
12823       531.8       531.8      186.55    194.25       412.60
11744       528.3       528.3      255.05    286.80       411.10
13217       528.2       528.2      252.75    286.75       409.65
16737       527.7       527.7      247.90    283.20       407.70
12483       527.6       527.6      254.65    286.40       410.75
12486       506.0       506.0      240.60    271.85       387.60
11748       506.0       506.0      240.60    271.85       387.60
11747       506.0       506.0      240.60    271.85       387.60
14092       506.0       506.0      240.60    271.85       387.60
16401       411.3       411.3      248.25    282.80       354.25
13758       407.4       407.4      245.50    279.75       350.70
16295       401.4       401.4      220.15    228.40       316.40
14483       400.5       400.5      219.65    228.25       315.45
13657       397.6       397.6      215.20    225.70       313.15
16469       379.5       379.5      207.20    229.65       295.60
13827       375.7       375.7      204.75    227.00       292.45
14655       375.1       375.1      204.70    226.90       292.15
14191       374.7       374.7      219.50    251.55       319.85
15314       372.7       372.7      218.10    250.05       318.05
16061       372.5       372.5      203.35    225.15       290.00
12588       372.3       372.3      217.20    249.20       317.45
15708       371.3       371.3      217.45    249.10       316.85
14923       370.9       370.9      216.70    248.55       316.35
17342       369.2       369.2      200.60    222.30       286.95
```

[9]:
```python
#make a function to compare a specific job in a specific province
def compare_job_province(job, province):
    df_filtered = df[(df['province'] == province) & (df['job'] == job)]
```

```
    df_grouped = df_filtered.groupby(['job', 'region'])['mid_mean'].mean().
 ↪reset_index()
    df_sorted = df_grouped.sort_values('mid_mean', ascending=False)
    return df_sorted
```

[10]:
```
# "ontario", "québec", "british columbia", "alberta", "manitoba",␣
 ↪"saskatchewan", "nova scotia", "new brunswick", "newfoundland & labrador",␣
 ↪"prince edward island"
df_job_province = compare_job_province('construction project coordinator',␣
 ↪'québec')
df_job_province
```

[10]:
```
                                job                          region  \
1   construction project coordinator             chaudière-appalaches
6   construction project coordinator                        montréal
7   construction project coordinator      montérégie - agglomeration
4   construction project coordinator                           laval
11  construction project coordinator  montérégie - vaudreuil-soulanges
13  construction project coordinator                       outaouais
12  construction project coordinator                 national capital
5   construction project coordinator                        mauricie
3   construction project coordinator                      laurentides
9   construction project coordinator    montérégie - la haute-yamaska
10  construction project coordinator     montérégie - les maskoutains
0   construction project coordinator                 centre-du-québec
2   construction project coordinator                          estrie
8   construction project coordinator    montérégie - brome-missisquoi

    mid_mean
1     121.15
6     121.10
7     119.40
4     119.35
11    106.20
13    106.10
12    103.80
5     103.05
3     102.45
9     101.80
10    101.75
0     100.75
2     100.45
8     100.05
```

0.1 Across the country

[11]:
```
#function to compare a specific job in a all provinces
def compare_job_all_provinces(job):
    df_filtered = df[(df['job'] == job)]
    df_grouped = df_filtered.groupby(['job', 'province'])['mid_mean'].mean().
 ↪reset_index()
    df_sorted = df_grouped.sort_values('mid_mean', ascending=False)
    return df_sorted
```

[12]:
```
df_job_all_provinces = compare_job_all_provinces('construction project␣
 ↪coordinator')
df_job_all_provinces
```

[12]:

	job	province	mid_mean
0	construction project coordinator	alberta	117.600000
3	construction project coordinator	ontario	112.402632
1	construction project coordinator	british columbia	111.375000
5	construction project coordinator	saskatchewan	110.775000
4	construction project coordinator	québec	107.671429
2	construction project coordinator	manitoba	105.600000

0.1.1 Detailled breakdown of the salary by province

[13]:
```
# Get the unique provinces
provinces = df['province'].unique()

# Define the job to compare
job = 'construction project coordinator'

# Loop over the provinces and compare the job in each province
for province in provinces:
    # Filter the dataframe by province and job
    df_filtered = df[(df['province'] == province) & (df['job'] == job)]

    # Group the resulting dataframe by job and region, and calculate the mean␣
 ↪mid salary
    df_grouped = df_filtered.groupby(['job', 'region'])['mid_mean'].mean().
 ↪reset_index()

    # Sort the grouped dataframe by the mid salary in descending order
    df_sorted = df_grouped.sort_values('mid_mean', ascending=False)

    # Print the sorted dataframe with the province name
    print(f'{province.capitalize()}:')
    print(df_sorted)
    print()
```

Alberta:

	job	region	mid_mean
3	construction project coordinator	northern alberta	125.80
0	construction project coordinator	calgary metropolitan region	121.15
1	construction project coordinator	edmonton metropolitan region	115.40
2	construction project coordinator	fort st. john	113.55
4	construction project coordinator	red deer	112.10

British columbia:

	job	region	mid_mean
3	construction project coordinator	surrey/delta	120.65
0	construction project coordinator	fraser valley	108.95
2	construction project coordinator	prince george	108.55
1	construction project coordinator	kelowna	107.35

Manitoba:

	job	region	mid_mean
0	construction project coordinator	winnipeg metropolitan region	105.6

New brunswick:

Empty DataFrame
Columns: [job, region, mid_mean]
Index: []

Newfoundland & labrador:

Empty DataFrame
Columns: [job, region, mid_mean]
Index: []

Nova scotia:

Empty DataFrame
Columns: [job, region, mid_mean]
Index: []

Ontario:

	job	region	mid_mean
6	construction project coordinator	greater toronto area	130.60
18	construction project coordinator	york region	128.00
2	construction project coordinator	durham region	127.30
12	construction project coordinator	peel region	122.00
7	construction project coordinator	halton region	118.35
10	construction project coordinator	ottawa metropolitan region	113.85
4	construction project coordinator	greater hamilton area	109.80
5	construction project coordinator	greater sudbury area	108.65
16	construction project coordinator	wellington county	108.55
14	construction project coordinator	simcoe county	108.00
0	construction project coordinator	brant county	107.90
3	construction project coordinator	frontenac county	107.75

```
17  construction project coordinator        windsor-essex county    107.70
1   construction project coordinator                   cambridge    107.50
15  construction project coordinator            waterloo region    107.25
8   construction project coordinator               london area    106.65
11  construction project coordinator             oxford county    106.25
9   construction project coordinator            niagara region    104.80
13  construction project coordinator         peterborough county    104.75

Prince edward island:
Empty DataFrame
Columns: [job, region, mid_mean]
Index: []

Québec:
                                job                            region  \
1   construction project coordinator             chaudière-appalaches
6   construction project coordinator                         montréal
7   construction project coordinator       montérégie - agglomeration
4   construction project coordinator                            laval
11  construction project coordinator  montérégie - vaudreuil-soulanges
13  construction project coordinator                        outaouais
12  construction project coordinator                 national capital
5   construction project coordinator                         mauricie
3   construction project coordinator                      laurentides
9   construction project coordinator     montérégie - la haute-yamaska
10  construction project coordinator      montérégie - les maskoutains
0   construction project coordinator                centre-du-québec
2   construction project coordinator                           estrie
8   construction project coordinator     montérégie - brome-missisquoi

    mid_mean
1     121.15
6     121.10
7     119.40
4     119.35
11    106.20
13    106.10
12    103.80
5     103.05
3     102.45
9     101.80
10    101.75
0     100.75
2     100.45
8     100.05

Saskatchewan:
                               job                      region  mid_mean
```

```
0  construction project coordinator     regina metropolitan area    111.20
1  construction project coordinator  saskatoon metropolitan area    110.35
```

1 Part 2

1.0.1 Let's find the job that has the biggest difference across the country

[14]:
```
#function to find the job with the largest difference, the user can choose between entry, mid or senior
def find_job_with_biggest_diff(level):
    #if level is not entry, mid or senior, return an error
    if level not in ['entry_mean', 'mid_mean', 'senior_mean']:
        return 'Error: level must be entry_mean, mid_mean or senior_mean'
    # Group by the 'job' column and find the difference between the maximum and minimum values of the 'mid' column for each group
    job_diffs = df.groupby('job')[[level]].apply(lambda x: (x.max() - x.min()).sum())

    # Find the job with the largest overall salary difference
    max_diff_job = job_diffs.idxmax()

    # Print the result
    print("The job with the largest overall salary difference is: {}".format(max_diff_job))
```

[15]:
```
find_job_with_biggest_diff('entry_mean')
```

```
The job with the largest overall salary difference is: head of workshop and assembly
```

[16]:
```
job_diffs = df.groupby('job')[["mid_mean"]].apply(lambda x: (x.max() - x.min()).sum())
job_diffs_df = job_diffs.to_frame('Difference')
```

[17]:
```
#sort the dataframe by the difference
job_diffs_df_sorted = job_diffs_df.sort_values('Difference', ascending=False)
job_diffs_df_sorted.head(30)
```

[17]:

job	Difference
head of workshop and assembly	206.50
chief financial officer	146.15
controller	97.00
project coordinator	90.90
general manager	85.35
contracts administrator	84.00

```
vice president finance                          79.95
data entry clerk                                74.00
manufacturing director                          73.65
inventory control manager                       70.65
parts manager                                   68.95
program manager                                 62.55
director of operations                          59.60
site supervisor                                 58.60
property manager                                52.55
transportation manager                          51.80
materials manager                               51.10
director - financial planning & analysis        50.15
finance director                                49.45
qa/qc inspector                                 48.95
claims supervisor                               48.20
office/operations manager                       47.55
executive assistant                             45.95
administrative manager                          44.20
vice-president internal audit                   43.25
health and safety manager                       42.55
vice president human resources / chro           42.35
bilingual customer service team lead            41.65
hr manager                                      41.40
planner/scheduler                               40.85
```

[ ]: