isabellaparlato.final.project

December 13, 2024

1 FINAL PROJECT

1.0.1 Isabella Parlato

1.0.2 Date: December 13, 2024

1.1 Prompt

You are a data scientist and would like to know where the top 5 places in the world (country or city) where your salary (in USD) will go the farthest with respect to each individual index within the cost_of_living.csv file. Provide a simple statistical analysis in a Jupyter Notebook file and provide visualizations to support your analysis (I am looking for data wrangling more than anything).

There are several ways to convert currencies to USD. Here are some examples in Python: https://pytutorial.com/currency-conversion-in-python

[1]:
```python
# import libraries

import numpy as np
import pandas as pd
import string as str
import matplotlib.pyplot as plt
import seaborn as sns
```

[2]:
```python
# confirm working directory and change to project folder with data

import os
os.chdir(r"C:\Users\bells\OneDrive\Documents\MS Data Sci GRAD␣
  ↪SCHOOL\DSE5002_R-Python\Project_2_python\data")

import os
print(os.getcwd())
```

```
C:\Users\bells\OneDrive\Documents\MS Data Sci GRAD SCHOOL\DSE5002_R-
Python\Project_2_python\data
```

1.2 Add In Data Sets & Clean Them Up

[3]:
```
top_countries_women = pd.read_excel('best_countries_women_2024.xlsx')
print(top_countries_women.head())

## I pulled this data from https://ceoworld.biz/2024/04/15/
↪revealed-worlds-best-countries-for-women-2024/
```

```
   Country Rankings as Best for Women      Country  \
0                                   1  Netherlands
1                                   2       Norway
2                                   3       Sweden
3                                   4      Denmark
4                                   5      Finland

   Country Score as Best for Women
0                             99.7
1                             99.4
2                             99.2
3                             98.7
4                             98.3
```

[4]:
```
cost_of_living = pd.read_csv('cost_of_living.csv')
print (cost_of_living.head())
```

```
   Rank                 City  Cost of Living Index  Rent Index  \
0   NaN     Hamilton, Bermuda               149.02       96.10
1   NaN   Zurich, Switzerland               131.24       69.26
2   NaN    Basel, Switzerland               130.93       49.38
3   NaN      Zug, Switzerland               128.13       72.12
4   NaN   Lugano, Switzerland               123.99       44.99

   Cost of Living Plus Rent Index  Groceries Index  Restaurant Price Index  \
0                          124.22           157.89                  155.22
1                          102.19           136.14                  132.52
2                           92.70           137.07                  130.95
3                          101.87           132.61                  130.93
4                           86.96           129.17                  119.80

   Local Purchasing Power Index
0                         79.43
1                        129.79
2                        111.53
3                        143.40
4                        111.96
```

[5]:
```
## cleaning up cost of living df - I want to take only the Countries from the␣
↪City column by just taking last word
```

```
cost_of_living['Country'] = cost_of_living['City'].str.split(',').str[-1]

#then remove any spaces so it will line up with out data set's Country columns

cost_of_living['Country'] = cost_of_living['Country'].str.strip()

print(cost_of_living.head())
print(cost_of_living.columns)
```

```
   Rank                 City  Cost of Living Index  Rent Index  \
0   NaN    Hamilton, Bermuda                149.02       96.10
1   NaN  Zurich, Switzerland                131.24       69.26
2   NaN   Basel, Switzerland                130.93       49.38
3   NaN     Zug, Switzerland                128.13       72.12
4   NaN  Lugano, Switzerland                123.99       44.99

   Cost of Living Plus Rent Index  Groceries Index  Restaurant Price Index  \
0                          124.22           157.89                  155.22
1                          102.19           136.14                  132.52
2                           92.70           137.07                  130.95
3                          101.87           132.61                  130.93
4                           86.96           129.17                  119.80

   Local Purchasing Power Index      Country
0                         79.43      Bermuda
1                        129.79  Switzerland
2                        111.53  Switzerland
3                        143.40  Switzerland
4                        111.96  Switzerland
Index(['Rank', 'City', 'Cost of Living Index', 'Rent Index',
       'Cost of Living Plus Rent Index', 'Groceries Index',
       'Restaurant Price Index', 'Local Purchasing Power Index', 'Country'],
      dtype='object')
```

[6]:
```
## now to delete unnecessary columns in cost of living df

cost_of_living = cost_of_living.drop(columns=['City','Rank'])
print(cost_of_living.head())
```

```
   Cost of Living Index  Rent Index  Cost of Living Plus Rent Index  \
0                149.02       96.10                          124.22
1                131.24       69.26                          102.19
2                130.93       49.38                           92.70
3                128.13       72.12                          101.87
4                123.99       44.99                           86.96

   Groceries Index  Restaurant Price Index  Local Purchasing Power Index  \
0          157.89                  155.22                         79.43
1          136.14                  132.52                        129.79
2          137.07                  130.95                        111.53
3          132.61                  130.93                        143.40
4          129.17                  119.80                        111.96

       Country
0      Bermuda
1  Switzerland
2  Switzerland
3  Switzerland
4  Switzerland
```

[7]:
```
country_codes = pd.read_excel('country_codes.xlsx')
print(country_codes.head())
print(country_codes.columns)
```

```
          Country Alpha-2 code Alpha-3 code  Numeric
0     Afghanistan           AF          AFG        4
1         Albania           AL          ALB        8
2         Algeria           DZ          DZA       12
3  American Samoa           AS          ASM       16
4         Andorra           AD          AND       20
Index(['Country', 'Alpha-2 code', 'Alpha-3 code', 'Numeric'], dtype='object')
```

[8]:
```
## now to clean up the country codes df

country_codes = country_codes.drop(columns=['Numeric'])
country_codes.rename(columns={'Alpha-2 code': 'Country Code_2 Letters',␣
 ↪'Alpha-3 code': 'Country Code_3 Letters'}, inplace=True)
print(country_codes.head())
```

```
          Country Country Code_2 Letters Country Code_3 Letters
0     Afghanistan                     AF                    AFG
1         Albania                     AL                    ALB
2         Algeria                     DZ                    DZA
3  American Samoa                     AS                    ASM
4         Andorra                     AD                    AND
```

[9]:
```
ds_salaries = pd.read_csv('ds_salaries.csv')
print(ds_salaries.head())
print(ds_salaries.columns)
```

```
   Unnamed: 0  work_year experience_level employment_type  \
0           0       2020               MI              FT
1           1       2020               SE              FT
2           2       2020               SE              FT
3           3       2020               MI              FT
4           4       2020               SE              FT
```

```
                    job_title  salary salary_currency  salary_in_usd  \
0              Data Scientist   70000             EUR          79833
1  Machine Learning Scientist  260000             USD         260000
2          Big Data Engineer   85000             GBP         109024
3        Product Data Analyst   20000             USD          20000
4   Machine Learning Engineer  150000             USD         150000

  employee_residence  remote_ratio company_location company_size
0                 DE             0               DE            L
1                 JP             0               JP            S
2                 GB            50               GB            M
3                 HN             0               HN            S
4                 US            50               US            L
Index(['Unnamed: 0', 'work_year', 'experience_level', 'employment_type',
       'job_title', 'salary', 'salary_currency', 'salary_in_usd',
       'employee_residence', 'remote_ratio', 'company_location',
       'company_size'],
      dtype='object')
```

[10]:
```
## now to clean up ds_salaries df

ds_salaries.drop(columns=['Unnamed: 0','work_year', 'experience_level',␣
 ↪'employment_type','salary', 'salary_currency','remote_ratio',␣
 ↪'company_location',
        'company_size'], inplace=True)
print(ds_salaries.columns)
```

```
Index(['job_title', 'salary_in_usd', 'employee_residence'], dtype='object')
```

[11]:
```
## now that it's down to 3 columns, I want to only keep the data science roles␣
 ↪then delete job_titles

ds_salaries = ds_salaries[ds_salaries['job_title'] == 'Data Scientist']
ds_salaries.drop(columns=['job_title'], inplace=True)
print(ds_salaries.head())
```

```
    salary_in_usd employee_residence
0           79833                 DE
7           35735                 HU
10          51321                 FR
11          40481                 IN
12          39916                 FR
```

[12]:
```
levels_fyi_salaries = pd.read_csv('levels_fyi_salary_data.csv')
print(levels_fyi_salaries.head())
print(levels_fyi_salaries.columns)
```

```
            timestamp    company level                         title  \
0   6/7/2017 11:33:27     Oracle    L3             Product Manager
1  6/10/2017 17:11:29       eBay  SE 2           Software Engineer
2  6/11/2017 14:53:57     Amazon    L7             Product Manager
3   6/17/2017 0:23:14      Apple    M1  Software Engineering Manager
4  6/20/2017 10:58:51  Microsoft    60           Software Engineer

   totalyearlycompensation          location  yearsofexperience  \
0                   127000   Redwood City, CA                1.5
1                   100000  San Francisco, CA                5.0
2                   310000        Seattle, WA                8.0
3                   372000      Sunnyvale, CA                7.0
4                   157000  Mountain View, CA                5.0

   yearsatcompany  tag  basesalary  …  Doctorate_Degree  Highschool  \
0             1.5  NaN    107000.0  …                 0           0
1             3.0  NaN         0.0  …                 0           0
2             0.0  NaN    155000.0  …                 0           0
3             5.0  NaN    157000.0  …                 0           0
4             3.0  NaN         0.0  …                 0           0

  Some_College Race_Asian  Race_White  Race_Two_Or_More  Race_Black  \
0            0          0           0                 0           0
1            0          0           0                 0           0
2            0          0           0                 0           0
3            0          0           0                 0           0
4            0          0           0                 0           0

   Race_Hispanic  Race  Education
0              0   NaN        NaN
1              0   NaN        NaN
2              0   NaN        NaN
3              0   NaN        NaN
4              0   NaN        NaN

[5 rows x 29 columns]
Index(['timestamp', 'company', 'level', 'title', 'totalyearlycompensation',
       'location', 'yearsofexperience', 'yearsatcompany', 'tag', 'basesalary',
       'stockgrantvalue', 'bonus', 'gender', 'otherdetails', 'cityid', 'dmaid',
       'rowNumber', 'Masters_Degree', 'Bachelors_Degree', 'Doctorate_Degree',
       'Highschool', 'Some_College', 'Race_Asian', 'Race_White',
       'Race_Two_Or_More', 'Race_Black', 'Race_Hispanic', 'Race', 'Education'],
      dtype='object')
```

[13]:
```
## now I want to clean up levels_fyi_salaries to remove some of the unnecessary␣
↪columns
```

```
levels_fyi_salaries.drop(columns=['timestamp', 'company',␣
 ↪'level','yearsofexperience', 'yearsatcompany', 'tag', 'gender',␣
 ↪'otherdetails', 'cityid', 'dmaid',
       'rowNumber', 'Masters_Degree', 'Bachelors_Degree', 'Doctorate_Degree',␣
 ↪'Highschool', 'Some_College', 'Race_Asian', 'Race_White', 'Race_Two_Or_More',
       'Race_Black', 'Race_Hispanic', 'Race', 'Education'], inplace=True)
print(levels_fyi_salaries.head())
```

```
                      title  totalyearlycompensation           location  \
0           Product Manager                   127000   Redwood City, CA
1         Software Engineer                   100000  San Francisco, CA
2           Product Manager                   310000        Seattle, WA
3  Software Engineering Manager               372000      Sunnyvale, CA
4         Software Engineer                   157000  Mountain View, CA

   basesalary  stockgrantvalue    bonus
0    107000.0          20000.0  10000.0
1         0.0              0.0      0.0
2    155000.0              0.0      0.0
3    157000.0         180000.0  35000.0
4         0.0              0.0      0.0
```

[14]:
```
## want to just narrow levels_fyi_salaries down to data scientists only

levels_fyi_salaries = levels_fyi_salaries[levels_fyi_salaries['title'] == 'Data␣
 ↪Scientist']
levels_fyi_salaries.drop(columns=['title'], inplace=True)
print(levels_fyi_salaries.head())
```

```
     totalyearlycompensation           location  basesalary  stockgrantvalue  \
419                   233000  San Francisco, CA    162000.0         220000.0
440                   218000        Seattle, WA    165000.0          28000.0
444                   180000       San Jose, CA         0.0              0.0
454                   500000  San Francisco, CA    200000.0         280000.0
495                   370000        Seattle, WA    190000.0         140000.0

       bonus
419  10000.0
440  23000.0
444      0.0
454  20000.0
495  40000.0
```

[15]:
```
## seems like for location, US only has city, state while other countries have␣
 ↪it after 3rd comma so want to get down to only countries column


## step 1 - separate into 3 columns by comma
```

```
levels_fyi_salaries[['City', 'State/County', 'Country']] =␣
 ↪levels_fyi_salaries['location'].str.split(',', expand=True)

## step 2 - clear up any extra spaces that may have pulled
levels_fyi_salaries['City'] = levels_fyi_salaries['City'].str.strip()
levels_fyi_salaries['State/County'] = levels_fyi_salaries['State/County'].str.
 ↪strip()
levels_fyi_salaries['Country'] = levels_fyi_salaries['Country'].str.strip()

## step 3 - since the US rows will pull blank to 3rd row, want to replace blank␣
 ↪with US
levels_fyi_salaries['Country'] = levels_fyi_salaries['Country'].fillna('United␣
 ↪States')

## step 4 - this should account for any missing data after filling above so␣
 ↪want to set this to US as well just in case
levels_fyi_salaries['Country'] = levels_fyi_salaries['Country'].replace('',␣
 ↪'United States')

## step 5 - now to drop the other unnecessary columns to leave Country only
levels_fyi_salaries.drop(columns=['City', 'State/County', 'location'],␣
 ↪inplace=True)

## finally print to see if it worked
print(levels_fyi_salaries)
```

```
       totalyearlycompensation  basesalary  stockgrantvalue    bonus  \
419                     233000    162000.0         220000.0  10000.0
440                     218000    165000.0          28000.0  23000.0
444                     180000         0.0              0.0      0.0
454                     500000    200000.0         280000.0  20000.0
495                     370000    190000.0         140000.0  40000.0
…                          …         …               …        …
62240                   155000    141000.0              0.0  14000.0
62283                   150000    150000.0          30000.0  30000.0
62285                   185000    150000.0          20000.0  15000.0
62529                   685000    221000.0         296000.0  55000.0
62623                   175000    135000.0          29000.0  11000.0

             Country
419    United States
440    United States
444    United States
454    United States
495    United States
…                …
62240  United States
```

```
62283      Singapore
62285  United States
62529  United States
62623  United States

[2578 rows x 5 columns]
```

1.3 Merging Data Sets

[16]:
```
## In order to figure out how to join them, I want to see what columns they␣
↪have in common

print('This is for the top_countries_women data set', top_countries_women.
↪columns)
print(top_countries_women.head())

print('This is for the cost_of_living data set', cost_of_living.columns)
print(cost_of_living.head())

print('This is for the country_codes data set', country_codes.columns)
print(country_codes.head())

print('This is for the ds_salaries data set', ds_salaries.columns)
print(ds_salaries.head())

print('This is for the levels_fyi_salaries data set', levels_fyi_salaries.
↪columns)
print(levels_fyi_salaries.head())
```

```
This is for the top_countries_women data set Index(['Country Rankings as Best
for Women', 'Country',
       'Country Score as Best for Women'],
      dtype='object')
   Country Rankings as Best for Women      Country  \
0                                    1  Netherlands
1                                    2       Norway
2                                    3       Sweden
3                                    4      Denmark
4                                    5      Finland

   Country Score as Best for Women
0                             99.7
1                             99.4
2                             99.2
3                             98.7
4                             98.3
This is for the cost_of_living data set Index(['Cost of Living Index', 'Rent
Index', 'Cost of Living Plus Rent Index',
       'Groceries Index', 'Restaurant Price Index',
       'Local Purchasing Power Index', 'Country'],
      dtype='object')
   Cost of Living Index  Rent Index  Cost of Living Plus Rent Index  \
0                149.02       96.10                           124.22
1                131.24       69.26                           102.19
2                130.93       49.38                            92.70
3                128.13       72.12                           101.87
4                123.99       44.99                            86.96

   Groceries Index  Restaurant Price Index  Local Purchasing Power Index  \
0           157.89                  155.22                         79.43
1           136.14                  132.52                        129.79
2           137.07                  130.95                        111.53
3           132.61                  130.93                        143.40
4           129.17                  119.80                        111.96

       Country
0      Bermuda
1  Switzerland
2  Switzerland
3  Switzerland
4  Switzerland
This is for the country_codes data set Index(['Country', 'Country Code_2
Letters', 'Country Code_3 Letters'], dtype='object')
          Country Country Code_2 Letters Country Code_3 Letters
0     Afghanistan                     AF                    AFG
1         Albania                     AL                    ALB
2         Algeria                     DZ                    DZA
3  American Samoa                     AS                    ASM
4         Andorra                     AD                    AND
This is for the ds_salaries data set Index(['salary_in_usd',
'employee_residence'], dtype='object')
    salary_in_usd employee_residence
0           79833                 DE
7           35735                 HU
10          51321                 FR
11          40481                 IN
12          39916                 FR
This is for the levels_fyi_salaries data set Index(['totalyearlycompensation',
'basesalary', 'stockgrantvalue', 'bonus',
       'Country'],
      dtype='object')
     totalyearlycompensation  basesalary  stockgrantvalue    bonus  \
419                   233000    162000.0         220000.0  10000.0
440                   218000    165000.0          28000.0  23000.0
444                   180000         0.0              0.0      0.0
454                   500000    200000.0         280000.0  20000.0
```

```
495                   370000    190000.0         140000.0  40000.0

           Country
419  United States
440  United States
444  United States
454  United States
495  United States
```

```
[17]: ## full join for top_countries_women AND cost_of_living to create: ␣
       ↪costs_top_countries_women

      costs_top_countries_women = pd.merge(cost_of_living, top_countries_women, how=␣
       ↪'outer', on= 'Country')
      print(costs_top_countries_women)
      print(costs_top_countries_women.columns)
```

```
     Cost of Living Index  Rent Index  Cost of Living Plus Rent Index  \
0                   21.35        3.17                           12.83
1                   38.68       11.33                           25.86
2                   29.84        6.67                           18.98
3                     NaN         NaN                             NaN
4                     NaN         NaN                             NaN
..                    …           …                             …
621                 39.01       17.85                           29.09
622                 36.85       12.21                           25.30
623                   NaN         NaN                             NaN
624                 33.57       10.18                           22.60
625                 45.69        9.56                           28.75

     Groceries Index  Restaurant Price Index  Local Purchasing Power Index  \
0              15.22                   14.85                         22.79
1              30.99                   29.86                         31.15
2              30.25                   20.79                         21.78
3                NaN                     NaN                           NaN
4                NaN                     NaN                           NaN
..               …                     …                           …
621            39.91                   21.57                         30.87
622            38.66                   19.21                         30.67
623              NaN                     NaN                           NaN
624            32.85                   23.63                         37.48
625            37.05                   39.05                         17.59

         Country  Country Rankings as Best for Women  \
0    Afghanistan                               149.0
1        Albania                                44.0
2        Algeria                               104.0
3        Andorra                                30.0
```

```
4         Angola                              148.0
..           …                                 …
621      Vietnam                               88.0
622      Vietnam                               88.0
623        Yemen                              147.0
624       Zambia                              121.0
625     Zimbabwe                              119.0

     Country Score as Best for Women
0                              41.60
1                              82.87
2                              72.74
3                              86.60
4                              42.25
..                               …
621                            75.52
622                            75.52
623                            43.95
624                            68.99
625                            69.43

[626 rows x 9 columns]
Index(['Cost of Living Index', 'Rent Index', 'Cost of Living Plus Rent Index',
       'Groceries Index', 'Restaurant Price Index',
       'Local Purchasing Power Index', 'Country',
       'Country Rankings as Best for Women',
       'Country Score as Best for Women'],
      dtype='object')
```

[18]:
```
## full join for costs_top_countries_women & country_codes to create: ␣
 ↪costs_top_countries_women_codes

costs_top_countries_women_codes = pd.merge(costs_top_countries_women,␣
 ↪country_codes, how= 'outer', on= 'Country')
print(costs_top_countries_women_codes)
print(costs_top_countries_women_codes.columns)
```

```
     Cost of Living Index  Rent Index  Cost of Living Plus Rent Index  \
0                   21.35        3.17                           12.83
1                   38.68       11.33                           25.86
2                   29.84        6.67                           18.98
3                     NaN         NaN                             NaN
4                     NaN         NaN                             NaN
..                    …          …                             …
726                   NaN         NaN                             NaN
727                   NaN         NaN                             NaN
728                 33.57       10.18                           22.60
729                 45.69        9.56                           28.75
```

```
730                  NaN        NaN                             NaN

     Groceries Index  Restaurant Price Index  Local Purchasing Power Index  \
0              15.22                   14.85                         22.79
1              30.99                   29.86                         31.15
2              30.25                   20.79                         21.78
3                NaN                     NaN                           NaN
4                NaN                     NaN                           NaN
..               …                       …                            …
726              NaN                     NaN                           NaN
727              NaN                     NaN                           NaN
728            32.85                   23.63                         37.48
729            37.05                   39.05                         17.59
730              NaN                     NaN                           NaN

            Country  Country Rankings as Best for Women  \
0       Afghanistan                               149.0
1           Albania                                44.0
2           Algeria                               104.0
3    American Samoa                                 NaN
4           Andorra                                30.0
..              …                                 …
726  Western Sahara                                 NaN
727           Yemen                               147.0
728          Zambia                               121.0
729        Zimbabwe                               119.0
730   Åland Islands                                 NaN

     Country Score as Best for Women Country Code_2 Letters  \
0                              41.60                     AF
1                              82.87                     AL
2                              72.74                     DZ
3                                NaN                     AS
4                              86.60                     AD
..                               …                    …
726                              NaN                     EH
727                            43.95                     YE
728                            68.99                     ZM
729                            69.43                     ZW
730                              NaN                     AX

    Country Code_3 Letters
0                      AFG
1                      ALB
2                      DZA
3                      ASM
4                      AND
..                     …
```

```
726                    ESH
727                    YEM
728                    ZMB
729                    ZWE
730                    ALA

[731 rows x 11 columns]
Index(['Cost of Living Index', 'Rent Index', 'Cost of Living Plus Rent Index',
       'Groceries Index', 'Restaurant Price Index',
       'Local Purchasing Power Index', 'Country',
       'Country Rankings as Best for Women', 'Country Score as Best for Women',
       'Country Code_2 Letters', 'Country Code_3 Letters'],
      dtype='object')
```

[19]:
```
## full join for costs_top_countries_women_codes & levels_fyi_salaries to␣
↪create:  women_costs_codes_levels

women_costs_codes_levels = pd.merge(costs_top_countries_women_codes,␣
↪levels_fyi_salaries, how= 'outer', on= 'Country')
print(women_costs_codes_levels)
print(women_costs_codes_levels.columns)
```

```
        Cost of Living Index  Rent Index  Cost of Living Plus Rent Index  \
0                      21.35        3.17                           12.83
1                      38.68       11.33                           25.86
2                      29.84        6.67                           18.98
3                        NaN         NaN                             NaN
4                        NaN         NaN                             NaN
…                        …           …                               …
221014                   NaN         NaN                             NaN
221015                   NaN         NaN                             NaN
221016                 33.57       10.18                           22.60
221017                 45.69        9.56                           28.75
221018                   NaN         NaN                             NaN

        Groceries Index  Restaurant Price Index  Local Purchasing Power Index  \
0                 15.22                   14.85                         22.79
1                 30.99                   29.86                         31.15
2                 30.25                   20.79                         21.78
3                   NaN                     NaN                           NaN
4                   NaN                     NaN                           NaN
…                  …                       …                             …
221014              NaN                     NaN                           NaN
221015              NaN                     NaN                           NaN
221016            32.85                   23.63                         37.48
221017            37.05                   39.05                         17.59
221018              NaN                     NaN                           NaN
```

```
               Country  Country Rankings as Best for Women  \
0          Afghanistan                               149.0
1              Albania                                44.0
2              Algeria                               104.0
3       American Samoa                                 NaN
4              Andorra                                30.0
…                    …                                 …
221014  Western Sahara                                 NaN
221015           Yemen                               147.0
221016          Zambia                               121.0
221017        Zimbabwe                               119.0
221018   Åland Islands                                 NaN

        Country Score as Best for Women Country Code_2 Letters  \
0                                 41.60                     AF
1                                 82.87                     AL
2                                 72.74                     DZ
3                                   NaN                     AS
4                                 86.60                     AD
…                                 …                     …
221014                              NaN                     EH
221015                            43.95                     YE
221016                            68.99                     ZM
221017                            69.43                     ZW
221018                              NaN                     AX

       Country Code_3 Letters  totalyearlycompensation  basesalary  \
0                         AFG                      NaN         NaN
1                         ALB                      NaN         NaN
2                         DZA                      NaN         NaN
3                         ASM                      NaN         NaN
4                         AND                      NaN         NaN
…                       …                      …          …
221014                    ESH                      NaN         NaN
221015                    YEM                      NaN         NaN
221016                    ZMB                      NaN         NaN
221017                    ZWE                      NaN         NaN
221018                    ALA                      NaN         NaN

        stockgrantvalue  bonus
0                   NaN    NaN
1                   NaN    NaN
2                   NaN    NaN
3                   NaN    NaN
4                   NaN    NaN
…                 …    …
221014              NaN    NaN
221015              NaN    NaN
```

```
221016              NaN    NaN
221017              NaN    NaN
221018              NaN    NaN

[221019 rows x 15 columns]
Index(['Cost of Living Index', 'Rent Index', 'Cost of Living Plus Rent Index',
       'Groceries Index', 'Restaurant Price Index',
       'Local Purchasing Power Index', 'Country',
       'Country Rankings as Best for Women', 'Country Score as Best for Women',
       'Country Code_2 Letters', 'Country Code_3 Letters',
       'totalyearlycompensation', 'basesalary', 'stockgrantvalue', 'bonus'],
      dtype='object')
```

[20]:

```
## full join for women_costs_codes_levels & levels_fyi_salaries to create: ␣
 ↪women_costs_codes_levels_ds_salaries

women_costs_codes_levels_ds_salaries = pd.merge(women_costs_codes_levels,
                                                ds_salaries,
                                                how= 'outer',
                                                left_on= 'Country Code_2␣
 ↪Letters',
                                                right_on= 'employee_residence')
print(women_costs_codes_levels_ds_salaries)
print(women_costs_codes_levels_ds_salaries.columns)
```

```
        Cost of Living Index  Rent Index  Cost of Living Plus Rent Index  \
0                        NaN         NaN                             NaN
1                        NaN         NaN                             NaN
2                      21.35        3.17                           12.83
3                        NaN         NaN                             NaN
4                        NaN         NaN                             NaN
…                          …           …                               …
268618                 55.92       23.17                           40.56
268619                 55.92       23.17                           40.56
268620                 45.31       12.08                           29.73
268621                 39.01       17.85                           29.09
268622                 36.85       12.21                           25.30

        Groceries Index  Restaurant Price Index  Local Purchasing Power Index  \
0                   NaN                     NaN                           NaN
1                   NaN                     NaN                           NaN
2                 15.22                   14.85                         22.79
3                   NaN                     NaN                           NaN
4                   NaN                     NaN                           NaN
…                     …                       …                             …
268618            54.45                   48.18                        118.77
268619            54.45                   48.18                        118.77
268620            37.60                   48.60                         15.87
```

```
268621          39.91                   21.57                          30.87
268622          38.66                   19.21                          30.67

                         Country  Country Rankings as Best for Women  \
0                        Andorra                                30.0
1       United Arab Emirates (the)                               NaN
2                    Afghanistan                               149.0
3            Antigua and Barbuda                                72.0
4                       Anguilla                                 NaN
…                            …                             …
268618             United States                                20.0
268619             United States                                20.0
268620                 Venezuela                                81.0
268621                   Vietnam                                88.0
268622                   Vietnam                                88.0

        Country Score as Best for Women Country Code_2 Letters  \
0                                 86.60                     AD
1                                   NaN                     AE
2                                 41.60                     AF
3                                 77.94                     AG
4                                   NaN                     AI
…                                 …                   …
268618                            90.30                    NaN
268619                            90.30                    NaN
268620                            76.77                    NaN
268621                            75.52                    NaN
268622                            75.52                    NaN

       Country Code_3 Letters  totalyearlycompensation  basesalary  \
0                         AND                      NaN         NaN
1                         ARE                      NaN         NaN
2                         AFG                      NaN         NaN
3                         ATG                      NaN         NaN
4                         AIA                      NaN         NaN
…                       …                        …          …
268618                    NaN                 685000.0    221000.0
268619                    NaN                 175000.0    135000.0
268620                    NaN                      NaN         NaN
268621                    NaN                      NaN         NaN
268622                    NaN                      NaN         NaN

        stockgrantvalue    bonus  salary_in_usd employee_residence
0                   NaN      NaN            NaN                NaN
1                   NaN      NaN            NaN                NaN
2                   NaN      NaN            NaN                NaN
3                   NaN      NaN            NaN                NaN
4                   NaN      NaN            NaN                NaN
```

```
…                 …         …             …                 …
268618           296000.0   55000.0            NaN                 NaN
268619            29000.0   11000.0            NaN                 NaN
268620                NaN       NaN            NaN                 NaN
268621                NaN       NaN            NaN                 NaN
268622                NaN       NaN            NaN                 NaN

[268623 rows x 17 columns]
Index(['Cost of Living Index', 'Rent Index', 'Cost of Living Plus Rent Index',
       'Groceries Index', 'Restaurant Price Index',
       'Local Purchasing Power Index', 'Country',
       'Country Rankings as Best for Women', 'Country Score as Best for Women',
       'Country Code_2 Letters', 'Country Code_3 Letters',
       'totalyearlycompensation', 'basesalary', 'stockgrantvalue', 'bonus',
       'salary_in_usd', 'employee_residence'],
      dtype='object')
```

1.4 Cleaning up fully merged data set containing the 5 dataframes

[21]:
```
## now to pull out as csv to see what it looks like externally

women_costs_codes_levels_ds_salaries.
 ↪to_csv('women_costs_codes_levels_ds_salaries.csv', index=False)
```

[22]:
```
## now to filter since I would only plan on living in any of the 20 top␣
 ↪countries for women (the US is ranked 20)

## originally got an error since 'Country Rankings as Best for Women' was␣
 ↪stored as a string and not an int so need to convert it to filter first

women_costs_codes_levels_ds_salaries['Country Rankings as Best for Women'] = pd.
 ↪to_numeric(
    women_costs_codes_levels_ds_salaries['Country Rankings as Best for Women'],
    errors='coerce')

## using coerce should convert any non-convertible values back to NaN

top_20_cost_of_living_salaries =␣
 ↪women_costs_codes_levels_ds_salaries[women_costs_codes_levels_ds_salaries['Country␣
 ↪Rankings as Best for Women'] <= 20]
print(top_20_cost_of_living_salaries)
print(top_20_cost_of_living_salaries.columns)
```

```
        Cost of Living Index  Rent Index  Cost of Living Plus Rent Index  \
11                     77.32       30.14                           55.20
12                     74.59       22.54                           50.19
13                     73.55       26.48                           51.49
14                     70.53       32.15                           52.54
```

```
15                      68.36      34.66                              52.56
…                       …        …                               …
268615                  55.92      23.17                              40.56
268616                  55.92      23.17                              40.56
268617                  55.92      23.17                              40.56
268618                  55.92      23.17                              40.56
268619                  55.92      23.17                              40.56

        Groceries Index  Restaurant Price Index  Local Purchasing Power Index  \
11                66.32                   79.87                         68.01
12                70.69                   62.68                         79.03
13                67.58                   75.39                         78.35
14                66.99                   68.46                         77.81
15                61.71                   66.93                         79.56
…                 …                    …                             …
268615            54.45                   48.18                        118.77
268616            54.45                   48.18                        118.77
268617            54.45                   48.18                        118.77
268618            54.45                   48.18                        118.77
268619            54.45                   48.18                        118.77

              Country  Country Rankings as Best for Women  \
11            Austria                                12.0
12            Austria                                12.0
13            Austria                                12.0
14            Austria                                12.0
15            Austria                                12.0
…               …                                 …
268615  United States                                20.0
268616  United States                                20.0
268617  United States                                20.0
268618  United States                                20.0
268619  United States                                20.0

        Country Score as Best for Women Country Code_2 Letters  \
11                                 95.2                     AT
12                                 95.2                     AT
13                                 95.2                     AT
14                                 95.2                     AT
15                                 95.2                     AT
…                                 …                     …
268615                             90.3                    NaN
268616                             90.3                    NaN
268617                             90.3                    NaN
268618                             90.3                    NaN
268619                             90.3                    NaN

       Country Code_3 Letters  totalyearlycompensation  basesalary  \
11                          AUT                    17000.0      16000.0
12                          AUT                    17000.0      16000.0
13                          AUT                    17000.0      16000.0
14                          AUT                    17000.0      16000.0
15                          AUT                    17000.0      16000.0
…                         …                      …          …
268615                      NaN                   190000.0     160000.0
268616                      NaN                   155000.0     141000.0
268617                      NaN                   185000.0     150000.0
268618                      NaN                   685000.0     221000.0
268619                      NaN                   175000.0     135000.0

        stockgrantvalue    bonus  salary_in_usd employee_residence
11                  0.0   1000.0        91237.0                 AT
12                  0.0   1000.0        91237.0                 AT
13                  0.0   1000.0        91237.0                 AT
14                  0.0   1000.0        91237.0                 AT
15                  0.0   1000.0        91237.0                 AT
…                 …      …            …                 …
268615              0.0  30000.0            NaN                NaN
268616              0.0  14000.0            NaN                NaN
268617          20000.0  15000.0            NaN                NaN
268618         296000.0  55000.0            NaN                NaN
268619          29000.0  11000.0            NaN                NaN

[227319 rows x 17 columns]
Index(['Cost of Living Index', 'Rent Index', 'Cost of Living Plus Rent Index',
       'Groceries Index', 'Restaurant Price Index',
       'Local Purchasing Power Index', 'Country',
       'Country Rankings as Best for Women', 'Country Score as Best for Women',
       'Country Code_2 Letters', 'Country Code_3 Letters',
       'totalyearlycompensation', 'basesalary', 'stockgrantvalue', 'bonus',
       'salary_in_usd', 'employee_residence'],
      dtype='object')
```

[23]:
```
## now to drop some of the columns that I don't plan on using (since country␣
↪had 0 NaN, removing alpha codes)

top_20_cost_of_living_salaries = top_20_cost_of_living_salaries.
↪drop(columns=['Country Code_2 Letters', 'Country Code_3 Letters',␣
↪'stockgrantvalue', 'bonus', 'employee_residence'])
print(top_20_cost_of_living_salaries.columns)
```

```
Index(['Cost of Living Index', 'Rent Index', 'Cost of Living Plus Rent Index',
       'Groceries Index', 'Restaurant Price Index',
       'Local Purchasing Power Index', 'Country',
       'Country Rankings as Best for Women', 'Country Score as Best for Women',
       'totalyearlycompensation', 'basesalary', 'salary_in_usd'],
       dtype='object')
```

[24]:
```
## I want to use 'totalyearlycompensation' and/or 'salary_in_usd' and/or␣
↪'basesalary' to compare to costs so I want to combine them into 1 grabbing␣
↪max

top_20_cost_of_living_salaries['DS Salary'] =␣
↪top_20_cost_of_living_salaries[['salary_in_usd', 'totalyearlycompensation',␣
↪'basesalary']].max(axis=1)
print(top_20_cost_of_living_salaries)
print(top_20_cost_of_living_salaries.isna().sum())
```

```
        Cost of Living Index  Rent Index  Cost of Living Plus Rent Index  \
11                     77.32       30.14                           55.20
12                     74.59       22.54                           50.19
13                     73.55       26.48                           51.49
14                     70.53       32.15                           52.54
15                     68.36       34.66                           52.56
…                        …           …                              …
268615                 55.92       23.17                           40.56
268616                 55.92       23.17                           40.56
268617                 55.92       23.17                           40.56
268618                 55.92       23.17                           40.56
268619                 55.92       23.17                           40.56

        Groceries Index  Restaurant Price Index  Local Purchasing Power Index  \
11                66.32                   79.87                         68.01
12                70.69                   62.68                         79.03
13                67.58                   75.39                         78.35
14                66.99                   68.46                         77.81
15                61.71                   66.93                         79.56
…                   …                       …                             …
268615            54.45                   48.18                        118.77
268616            54.45                   48.18                        118.77
268617            54.45                   48.18                        118.77
268618            54.45                   48.18                        118.77
268619            54.45                   48.18                        118.77

              Country  Country Rankings as Best for Women  \
11            Austria                                12.0
12            Austria                                12.0
13            Austria                                12.0
14            Austria                                12.0
15            Austria                                12.0
…               …                                    …
268615  United States                                20.0
268616  United States                                20.0
268617  United States                                20.0
```

```
268618  United States                                        20.0
268619  United States                                        20.0

        Country Score as Best for Women  totalyearlycompensation  basesalary  \
11                                 95.2                  17000.0     16000.0
12                                 95.2                  17000.0     16000.0
13                                 95.2                  17000.0     16000.0
14                                 95.2                  17000.0     16000.0
15                                 95.2                  17000.0     16000.0
…                                   …                        …          …
268615                             90.3                 190000.0    160000.0
268616                             90.3                 155000.0    141000.0
268617                             90.3                 185000.0    150000.0
268618                             90.3                 685000.0    221000.0
268619                             90.3                 175000.0    135000.0

        salary_in_usd  DS Salary
11            91237.0    91237.0
12            91237.0    91237.0
13            91237.0    91237.0
14            91237.0    91237.0
15            91237.0    91237.0
…                 …          …
268615            NaN   190000.0
268616            NaN   155000.0
268617            NaN   185000.0
268618            NaN   685000.0
268619            NaN   175000.0

[227319 rows x 13 columns]
Cost of Living Index                        0
Rent Index                                  0
Cost of Living Plus Rent Index              0
Groceries Index                             0
Restaurant Price Index                      0
Local Purchasing Power Index                0
Country                                     0
Country Rankings as Best for Women          0
Country Score as Best for Women             0
totalyearlycompensation                    27
basesalary                                 27
salary_in_usd                          213518
DS Salary                                  27
dtype: int64
```

[25]:
```
## now to drop the 3 columns I combined into the new DS Salary Column
```

```
top_20_cost_of_living_salaries = top_20_cost_of_living_salaries.
 ↪drop(columns=['salary_in_usd', 'totalyearlycompensation', 'basesalary'])
print(top_20_cost_of_living_salaries.columns)
```

```
Index(['Cost of Living Index', 'Rent Index', 'Cost of Living Plus Rent Index',
       'Groceries Index', 'Restaurant Price Index',
       'Local Purchasing Power Index', 'Country',
       'Country Rankings as Best for Women', 'Country Score as Best for Women',
       'DS Salary'],
      dtype='object')
```

[26]:
```
top_20_cost_of_living_salaries.to_csv('top_20_cost_of_living_salaries.csv',␣
 ↪index=False)
```

1.5 Comparing DS Salary to Cost of Living Indexes

[27]:
```
## going to try to a loop to run through all 6 indexes in comparison to DS␣
 ↪Salary and create new columns for comparison

for indexes in [
    'Cost of Living Index'
    ,'Rent Index'
    ,'Cost of Living Plus Rent Index'
    ,'Groceries Index'
    ,'Restaurant Price Index'
    ,'Local Purchasing Power Index'
]:

    top_20_cost_of_living_salaries[f'How far Salary will go for {indexes}'] =␣
 ↪top_20_cost_of_living_salaries['DS Salary'] /␣
 ↪top_20_cost_of_living_salaries[indexes]

## colon after bracket to signify start of loop body

print(top_20_cost_of_living_salaries.head())

## by dividing salary by cost of living index, those resulting values should be␣
 ↪how many times your salary exceeds the value of the index
## basically results show number of times salary can pay for cost given by␣
 ↪index. higher number means salary is much greater than costs
```

```
    Cost of Living Index  Rent Index  Cost of Living Plus Rent Index  \
11                 77.32       30.14                           55.20
12                 74.59       22.54                           50.19
13                 73.55       26.48                           51.49
14                 70.53       32.15                           52.54
15                 68.36       34.66                           52.56
```

```
    Groceries Index  Restaurant Price Index  Local Purchasing Power Index  \
11            66.32                   79.87                         68.01
12            70.69                   62.68                         79.03
13            67.58                   75.39                         78.35
14            66.99                   68.46                         77.81
15            61.71                   66.93                         79.56

    Country  Country Rankings as Best for Women  \
11  Austria                                12.0
12  Austria                                12.0
13  Austria                                12.0
14  Austria                                12.0
15  Austria                                12.0

    Country Score as Best for Women  DS Salary  \
11                             95.2    91237.0
12                             95.2    91237.0
13                             95.2    91237.0
14                             95.2    91237.0
15                             95.2    91237.0

    How far Salary will go for Cost of Living Index  \
11                                      1179.992240
12                                      1223.180051
13                                      1240.475867
14                                      1293.591380
15                                      1334.654769

    How far Salary will go for Rent Index  \
11                            3027.106835
12                            4047.781721
13                            3445.506042
14                            2837.853810
15                            2632.342758

    How far Salary will go for Cost of Living Plus Rent Index  \
11                                        1652.844203
12                                        1817.832237
13                                        1771.936298
14                                        1736.524553
15                                        1735.863775

    How far Salary will go for Groceries Index  \
11                                 1375.708685
12                                 1290.663460
13                                 1350.059189
14                                 1361.949545
15                                 1478.479987
```

```
    How far Salary will go for Restaurant Price Index  \
11                                          1142.318768
12                                          1455.599872
13                                          1210.200292
14                                          1332.705229
15                                          1363.170477

    How far Salary will go for Local Purchasing Power Index
11                                          1341.523305
12                                          1154.460332
13                                          1164.479898
14                                          1172.561367
15                                          1146.769734
```

f-string (AKA f used above in body)

I wanted something so I didn't have to go through each index individually and type out a new column name every time (or copy and paste it) and then have to keep going through same function. This feature is supposed to simplify the process of doing that in strings. I guess other ways to do this same thing are modulo operator (%), which is older, or str.format() method. But since this f-string method seemed shorter (and newer) I figured I'd see if I could get it to work correctly

Resources: "f-strings in Python" by Geeks for Geeks - https://www.geeksforgeeks.org/formatted-string-literals-f-strings-python/ "Python's F-String for String Interpolation and Formatting" by Joanna Jablonski - https://realpython.com/python-f-strings/

[28]:
```
print(top_20_cost_of_living_salaries.sum())
```

```
Cost of Living Index
16610166.9
Rent Index
10344664.62
Cost of Living Plus Rent Index
13673117.3
Groceries Index
16672206.55
Restaurant Price Index
16400904.22
Local Purchasing Power Index
26299302.67
Country
AustriaAustriaAustriaAustriaAustriaAustraliaAu…
Country Rankings as Best for Women
4362511.0
Country Score as Best for Women
20622696.76
DS Salary
48325764640.0
How far Salary will go for Cost of Living Index
```

```
669450735.823406
How far Salary will go for Rent Index
1187807356.041467
How far Salary will go for Cost of Living Plus Rent Index
828958961.694786
How far Salary will go for Groceries Index
671110833.499928
How far Salary will go for Restaurant Price Index
679737965.328514
How far Salary will go for Local Purchasing Power Index
430075916.764321
dtype: object
```

[31]:
```
## okay so now i want to try using groupby to try to get country mean per cost␣
↪of living how far columns

country_means = top_20_cost_of_living_salaries.groupby('Country').
↪mean(numeric_only=True)

print(country_means)
print(country_means.columns)
```

```
                Cost of Living Index  Rent Index  \
Country
Australia                  77.601000   39.267000
Austria                    72.870000   29.194000
Canada                     71.959643   36.316786
Denmark                    84.295000   34.385000
Finland                    76.100000   28.778000
France                     77.741667   31.056667
Germany                    67.281154   28.933846
Italy                      67.687368   22.963158
Japan                      79.313333   43.150000
Luxembourg                 82.990000   63.430000
Netherlands                76.229286   37.622143
New Zealand                75.362000   36.112000
Norway                    101.148000   38.308000
Portugal                   48.445000   22.761250
Singapore                  83.980000   66.430000
Spain                      54.757857   22.733571
Sweden                     75.460000   31.642000
Switzerland               124.075714   59.495714
United Kingdom             70.495588   32.788235
United States              73.252000   46.378105

                Cost of Living Plus Rent Index  Groceries Index  \
Country
Australia                            59.633000        77.064000
```

```
Austria                             52.396000        66.658000
Canada                              55.252143        71.140714
Denmark                             60.905000        68.782500
Finland                             53.918000        68.764000
France                              55.858333        77.201667
Germany                             49.305385        52.963077
Italy                               46.722105        59.678947
Japan                               62.363333        86.570000
Luxembourg                          73.820000        75.830000
Netherlands                         58.132143        65.190000
New Zealand                         56.966000        74.206000
Norway                              71.694000        97.828000
Portugal                            36.407500        39.070000
Singapore                           75.750000        77.080000
Spain                               39.747143        46.930000
Sweden                              54.922000        68.756000
Switzerland                         93.802857       126.945714
United Kingdom                      52.819706        58.282059
United States                       60.654421        74.003684

                Restaurant Price Index  Local Purchasing Power Index  \
Country
Australia                    75.104000                    106.108000
Austria                      70.666000                     76.552000
Canada                       70.154643                     90.766429
Denmark                     100.470000                     95.470000
Finland                      85.516000                     92.624000
France                       75.265000                     82.336667
Germany                      65.510769                    100.196538
Italy                        68.949474                     59.398947
Japan                        45.650000                     75.850000
Luxembourg                   95.170000                    100.190000
Netherlands                  77.171429                     86.550000
New Zealand                  71.976000                     92.486000
Norway                      105.848000                     87.100000
Portugal                     43.557500                     46.712500
Singapore                    61.170000                     91.340000
Spain                        54.395000                     70.442857
Sweden                       77.452000                     94.060000
Switzerland                 126.914286                    122.392857
United Kingdom               76.595000                     90.622941
United States                72.332842                    117.364000

                Country Rankings as Best for Women  \
Country
Australia                                     16.0
Austria                                       12.0
Canada                                         6.0
```

```
Denmark                                   4.0
Finland                                   5.0
France                                    9.0
Germany                                  10.0
Italy                                    13.0
Japan                                    15.0
Luxembourg                               11.0
Netherlands                               1.0
New Zealand                               8.0
Norway                                    2.0
Portugal                                 18.0
Singapore                                19.0
Spain                                    14.0
Sweden                                    3.0
Switzerland                               7.0
United Kingdom                           17.0
United States                            20.0

                Country Score as Best for Women      DS Salary  \
Country
Australia                                 92.08  147940.600000
Austria                                   95.20   91237.000000
Canada                                    97.50  123333.216981
Denmark                                   98.70            NaN
Finland                                   98.30            NaN
France                                    96.40  171000.000000
Germany                                   95.90  110048.200000
Italy                                     94.80   44500.000000
Japan                                     93.69   65000.000000
Luxembourg                                95.70   94250.000000
Netherlands                               99.70  117428.571429
New Zealand                               96.80            NaN
Norway                                    99.40            NaN
Portugal                                  91.23            NaN
Singapore                                 90.68  143275.400000
Spain                                     94.40   70666.666667
Sweden                                    99.20   94166.666667
Switzerland                               97.10  171291.142857
United Kingdom                            91.26  156297.872340
United States                             90.30  219319.712618

                How far Salary will go for Cost of Living Index  \
Country
Australia                                           1908.979182
Austria                                             1254.378861
Canada                                              1717.629936
Denmark                                                     NaN
Finland                                                     NaN
```

```
France                                                 2209.398742
Germany                                                1640.866206
Italy                                                   661.004417
Japan                                                   822.055070
Luxembourg                                             1135.678997
Netherlands                                            1545.214903
New Zealand                                                    NaN
Norway                                                         NaN
Portugal                                                       NaN
Singapore                                              1706.065730
Spain                                                  1297.203842
Sweden                                                 1248.876177
Switzerland                                            1383.859392
United Kingdom                                         2227.379854
United States                                          3035.112762

                How far Salary will go for Rent Index  \
Country
Australia                                 3858.883515
Austria                                   3198.118233
Canada                                    3626.698800
Denmark                                           NaN
Finland                                           NaN
France                                    5958.034123
Germany                                   3987.018007
Italy                                     2068.229385
Japan                                     1627.181586
Luxembourg                                1485.889957
Netherlands                               3271.398009
New Zealand                                       NaN
Norway                                            NaN
Portugal                                          NaN
Singapore                                 2156.787596
Spain                                     3270.505468
Sweden                                    3075.229597
Switzerland                               2999.602330
United Kingdom                            5216.128600
United States                             5325.940836

                How far Salary will go for Cost of Living Plus Rent Index  \
Country
Australia                                            2491.586246
Austria                                              1743.000213
Canada                                               2249.980061
Denmark                                                      NaN
Finland                                                      NaN
France                                               3106.509741
Germany                                              2251.730084
```

```
Italy                                                        963.528713
Japan                                                       1051.516539
Luxembourg                                                  1276.754267
Netherlands                                                 2037.940677
New Zealand                                                         NaN
Norway                                                              NaN
Portugal                                                            NaN
Singapore                                                   1891.424422
Spain                                                       1797.136130
Sweden                                                      1722.434955
Switzerland                                                 1834.366997
United Kingdom                                              3009.589166
United States                                               3746.014561

                How far Salary will go for Groceries Index  \
Country
Australia                                        1923.367432
Austria                                          1371.372173
Canada                                           1744.426800
Denmark                                                  NaN
Finland                                                  NaN
France                                           2222.878979
Germany                                          2090.857905
Italy                                             754.784619
Japan                                             754.201658
Luxembourg                                       1242.911776
Netherlands                                      1802.823227
New Zealand                                              NaN
Norway                                                   NaN
Portugal                                                 NaN
Singapore                                        1858.788272
Spain                                            1515.002448
Sweden                                           1371.567062
Switzerland                                      1355.745836
United Kingdom                                   2698.295552
United States                                    3028.054294

                How far Salary will go for Restaurant Price Index  \
Country
Australia                                               1975.960833
Austria                                                 1300.798928
Canada                                                  1771.198495
Denmark                                                         NaN
Finland                                                         NaN
France                                                  2299.362352
Germany                                                 1691.956403
Italy                                                    651.633106
Japan                                                   1440.396902
```

```
Luxembourg                                                 990.333088
Netherlands                                               1532.825360
New Zealand                                                       NaN
Norway                                                            NaN
Portugal                                                          NaN
Singapore                                                 2342.249469
Spain                                                     1312.200766
Sweden                                                    1217.795813
Switzerland                                               1351.500333
United Kingdom                                            2058.544192
United States                                             3081.553754

                How far Salary will go for Local Purchasing Power Index
Country
Australia                                                 1397.302589
Austria                                                   1195.958927
Canada                                                    1411.351485
Denmark                                                           NaN
Finland                                                           NaN
France                                                    2108.806428
Germany                                                   1116.545066
Italy                                                      755.615086
Japan                                                      887.721789
Luxembourg                                                 940.712646
Netherlands                                               1364.800446
New Zealand                                                       NaN
Norway                                                            NaN
Portugal                                                          NaN
Singapore                                                 1568.594263
Spain                                                     1012.927047
Sweden                                                    1009.348064
Switzerland                                               1410.301976
United Kingdom                                            1746.655085
United States                                             1932.056143
Index(['Cost of Living Index', 'Rent Index', 'Cost of Living Plus Rent Index',
       'Groceries Index', 'Restaurant Price Index',
       'Local Purchasing Power Index', 'Country Rankings as Best for Women',
       'Country Score as Best for Women', 'DS Salary',
       'How far Salary will go for Cost of Living Index',
       'How far Salary will go for Rent Index',
       'How far Salary will go for Cost of Living Plus Rent Index',
       'How far Salary will go for Groceries Index',
       'How far Salary will go for Restaurant Price Index',
       'How far Salary will go for Local Purchasing Power Index'],
      dtype='object')
```

```
[32]: ## since that pulled the mean for all columns by County (now an index/row␣
       ↪header) going to drop some unneeded columns again

      country_means = country_means.drop(columns=['Cost of Living Index', 'Rent␣
       ↪Index', 'Cost of Living Plus Rent Index',
                                                  'Groceries Index', 'Restaurant␣
       ↪Price Index',
                                                  'Local Purchasing Power Index', 'DS␣
       ↪Salary'])
      print(country_means.columns)
```

```
Index(['Country Rankings as Best for Women', 'Country Score as Best for Women',
       'How far Salary will go for Cost of Living Index',
       'How far Salary will go for Rent Index',
       'How far Salary will go for Cost of Living Plus Rent Index',
       'How far Salary will go for Groceries Index',
       'How far Salary will go for Restaurant Price Index',
       'How far Salary will go for Local Purchasing Power Index'],
      dtype='object')
```

```
[34]: print(country_means)
```

```
                Country Rankings as Best for Women  \
Country
Australia                                     16.0
Austria                                       12.0
Canada                                         6.0
Denmark                                        4.0
Finland                                        5.0
France                                         9.0
Germany                                       10.0
Italy                                         13.0
Japan                                         15.0
Luxembourg                                    11.0
Netherlands                                    1.0
New Zealand                                    8.0
Norway                                         2.0
Portugal                                      18.0
Singapore                                     19.0
Spain                                         14.0
Sweden                                         3.0
Switzerland                                    7.0
United Kingdom                                17.0
United States                                 20.0

                Country Score as Best for Women  \
Country
Australia                                 92.08
```

```
Austria                                    95.20
Canada                                     97.50
Denmark                                    98.70
Finland                                    98.30
France                                     96.40
Germany                                    95.90
Italy                                      94.80
Japan                                      93.69
Luxembourg                                 95.70
Netherlands                                99.70
New Zealand                                96.80
Norway                                     99.40
Portugal                                   91.23
Singapore                                  90.68
Spain                                      94.40
Sweden                                     99.20
Switzerland                                97.10
United Kingdom                             91.26
United States                              90.30

                How far Salary will go for Cost of Living Index  \
Country
Australia                                             1908.979182
Austria                                               1254.378861
Canada                                                1717.629936
Denmark                                                       NaN
Finland                                                       NaN
France                                                2209.398742
Germany                                               1640.866206
Italy                                                  661.004417
Japan                                                  822.055070
Luxembourg                                            1135.678997
Netherlands                                           1545.214903
New Zealand                                                   NaN
Norway                                                        NaN
Portugal                                                      NaN
Singapore                                             1706.065730
Spain                                                 1297.203842
Sweden                                                1248.876177
Switzerland                                           1383.859392
United Kingdom                                        2227.379854
United States                                         3035.112762

                How far Salary will go for Rent Index  \
Country
Australia                                  3858.883515
Austria                                    3198.118233
Canada                                     3626.698800
```

```
Denmark                                    NaN
Finland                                    NaN
France                             5958.034123
Germany                            3987.018007
Italy                              2068.229385
Japan                              1627.181586
Luxembourg                         1485.889957
Netherlands                        3271.398009
New Zealand                                NaN
Norway                                     NaN
Portugal                                   NaN
Singapore                          2156.787596
Spain                              3270.505468
Sweden                             3075.229597
Switzerland                        2999.602330
United Kingdom                     5216.128600
United States                      5325.940836

                How far Salary will go for Cost of Living Plus Rent Index  \
Country
Australia                                                2491.586246
Austria                                                  1743.000213
Canada                                                   2249.980061
Denmark                                                          NaN
Finland                                                          NaN
France                                                   3106.509741
Germany                                                  2251.730084
Italy                                                     963.528713
Japan                                                    1051.516539
Luxembourg                                               1276.754267
Netherlands                                              2037.940677
New Zealand                                                      NaN
Norway                                                           NaN
Portugal                                                         NaN
Singapore                                                1891.424422
Spain                                                    1797.136130
Sweden                                                   1722.434955
Switzerland                                              1834.366997
United Kingdom                                           3009.589166
United States                                            3746.014561

                How far Salary will go for Groceries Index  \
Country
Australia                                         1923.367432
Austria                                           1371.372173
Canada                                            1744.426800
Denmark                                                   NaN
Finland                                                   NaN
```

```
France                                       2222.878979
Germany                                      2090.857905
Italy                                         754.784619
Japan                                         754.201658
Luxembourg                                   1242.911776
Netherlands                                  1802.823227
New Zealand                                          NaN
Norway                                               NaN
Portugal                                             NaN
Singapore                                    1858.788272
Spain                                        1515.002448
Sweden                                       1371.567062
Switzerland                                  1355.745836
United Kingdom                               2698.295552
United States                                3028.054294

                How far Salary will go for Restaurant Price Index  \
Country
Australia                                           1975.960833
Austria                                             1300.798928
Canada                                              1771.198495
Denmark                                                     NaN
Finland                                                     NaN
France                                              2299.362352
Germany                                             1691.956403
Italy                                                651.633106
Japan                                               1440.396902
Luxembourg                                           990.333088
Netherlands                                         1532.825360
New Zealand                                                 NaN
Norway                                                      NaN
Portugal                                                    NaN
Singapore                                           2342.249469
Spain                                               1312.200766
Sweden                                              1217.795813
Switzerland                                         1351.500333
United Kingdom                                      2058.544192
United States                                       3081.553754

                How far Salary will go for Local Purchasing Power Index
Country
Australia                                           1397.302589
Austria                                             1195.958927
Canada                                              1411.351485
Denmark                                                     NaN
Finland                                                     NaN
France                                              2108.806428
Germany                                             1116.545066
```

```
Italy                                                   755.615086
Japan                                                   887.721789
Luxembourg                                              940.712646
Netherlands                                            1364.800446
New Zealand                                                    NaN
Norway                                                         NaN
Portugal                                                       NaN
Singapore                                              1568.594263
Spain                                                  1012.927047
Sweden                                                 1009.348064
Switzerland                                            1410.301976
United Kingdom                                         1746.655085
United States                                          1932.056143
```

1.5.1 Summary of the Top 6 Countries in each Cost of Living Categories

Cost of Living Index 1. United States 2. United Kingdom 3. France 4. Australia 5. Canada 6. Singapore

Rent Index 1. France 2. United States 3. United Kingdom 4. Germany 5. Australia 6. Canada

Cost of Living Plus Rent Index 1. United States 2. France 3. United Kingdom 4. Australia 5. Germany 6. Canada

Groceries Index 1. United States 2. United Kingdom 3. France 4. Germany 5. Australia 6. Singapore

Restaurant Price Index 1. United States 2. Singapore 3. France 4. United Kingdom 5. Australia 6. Canada

Local Purchasing Power Index 1. France 2. United States 3. United Kingdom 4. Singapore 5. Canada 6. Switzerland

Reminder of Where the Countries that were named in the top 6 for each category place as Best for Women in 2024

Australia 16.0
Canada 6.0
France 9.0
Germany 10.0
Singapore 19.0
Switzerland 7.0
United Kingdom 17.0
United States 20.0

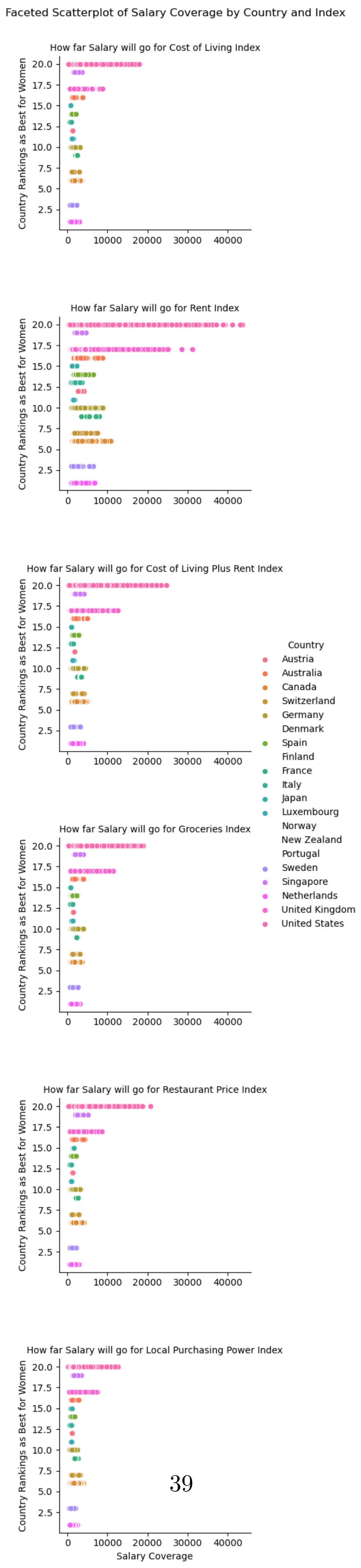
1.6 Visualizations

[67]:
```
# need to melt to long form since i think that will make ploting easier

top_20_cost_of_living_salaries_melt = top_20_cost_of_living_salaries.
 ↪melt(id_vars= ['Country', 'Country Rankings as Best for Women']
                                          ,value_vars= ['How far Salary will go␣
 ↪for Cost of Living Index'
                                                       ,'How far Salary will go␣
 ↪for Rent Index'
                                                       ,'How far Salary will go␣
 ↪for Cost of Living Plus Rent Index'
                                                       ,'How far Salary will go␣
 ↪for Groceries Index'
                                                       ,'How far Salary will go␣
 ↪for Restaurant Price Index'
                                                       ,'How far Salary will go␣
 ↪for Local Purchasing Power Index']
                                          ,var_name= 'Index'
                                          ,value_name= 'How Far Salary Covers')

print(top_20_cost_of_living_salaries_melt)
print(top_20_cost_of_living_salaries_melt.columns)
```

```
               Country  Country Rankings as Best for Women  \
0              Austria                                12.0
1              Austria                                12.0
2              Austria                                12.0
3              Austria                                12.0
4              Austria                                12.0
…                 …                               …
1363909  United States                                20.0
1363910  United States                                20.0
1363911  United States                                20.0
1363912  United States                                20.0
1363913  United States                                20.0

                                                  Index  \
0           How far Salary will go for Cost of Living Index
1           How far Salary will go for Cost of Living Index
2           How far Salary will go for Cost of Living Index
3           How far Salary will go for Cost of Living Index
4           How far Salary will go for Cost of Living Index
…                                                   …
1363909  How far Salary will go for Local Purchasing Po…
1363910  How far Salary will go for Local Purchasing Po…
1363911  How far Salary will go for Local Purchasing Po…
1363912  How far Salary will go for Local Purchasing Po…
1363913  How far Salary will go for Local Purchasing Po…

         How Far Salary Covers
0                  1179.992240
1                  1223.180051
```

```
2                 1240.475867
3                 1293.591380
4                 1334.654769
…                       …
1363909           1599.730572
1363910           1305.043361
1363911           1557.632399
1363912           5767.449693
1363913           1473.436053

[1363914 rows x 4 columns]
Index(['Country', 'Country Rankings as Best for Women', 'Index',
       'How Far Salary Covers'],
      dtype='object')
```

[52]:
```
## Scatterplot faceted

g = sns.FacetGrid(top_20_cost_of_living_salaries_melt, col='Index',␣
 ↪hue='Country', height=4, aspect=1, col_wrap=1)
g.map(sns.scatterplot, 'How Far Salary Covers', 'Country Rankings as Best for␣
 ↪Women')
g.add_legend()
g.set_titles(col_template='{col_name}')

g.fig.suptitle('Faceted Scatterplot of Salary Coverage by Country and Index')

# below should make it so the labels so on every facet and not just bottom one
g.set_axis_labels('Salary Coverage', 'Country Rankings as Best for Women')
g.tick_params(axis='x', labelbottom=True)

# below fixes space between stacked faceted plots
plt.subplots_adjust(hspace=0.5, top=0.95)

plt.show()
```


Faceted Scatterplot of Salary Coverage by Country and Index
How far Salary will go for Cost of Living Index
How far Salary will go for Rent Index
How far Salary will go for Cost of Living Plus Rent Index
How far Salary will go for Groceries Index
How far Salary will go for Restaurant Price Index
How far Salary will go for Local Purchasing Power Index
Country Rankings as Best for Women
Salary Coverage
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0
10000
20000
30000
40000
Country
Austria
Australia
Canada
Switzerland
Germany
Denmark
Spain
Finland
France
Italy
Japan
Luxembourg
Norway
New Zealand
Portugal
Sweden
Singapore
Netherlands
United Kingdom
United States

1.7 Final Thoughts/Analysis of my Top Five Countries

After reviewing the data and taking into account my personal preferences, my list of the five (5) countries would be as follows where I would consider moving to: 1. Canada 2. France 3. Australia 4. Singapore 5. United Kingdom

I already went into this knowing that I wouldn't plan on living in any country where it isn't great to be a women in that environment. So since the United States ranked #20 on that list of Best Countries for Women in 2024, I wouldn't plan on settling for less than what I'm used to. After reviewing the means by Country for each of the Cost of Living Indexes, I ended up with eight (8) countries that ended up being named as one of the top six (6) in at least one of the categories.

The United States seemingly did well in most categories but that might also be because of it likely having the most salary data available to work with so any outliers wouldn't have made too much of an impact. I left it off my list because I would love to travel more and living out of the United States for a bit is one of my goals. Switzerland only appeared in Local Purchasing Power Index as rank #6 so since it wasn't consistently in any of the other lists, I ruled it out from my list. I didn't put Germany on my list because recently I went on a trip to Costa Rica this fall and my roommate for the trip was a German woman, who was incredibly difficult to deal with. Thus, due to my recent poor experience, I'm not trying to go to Germany anytime soon.

I've been to Canada and Australia before and had great times in both. There's definitely a lot left of both countries that I would love to explore so they made my list. France did well in the means lists and I've heard good things about the amount of vacation time that employers give so that country ended up on my list as well. I think the United Kingdom would be fun to explore and an easy transition to live in given that English is the primary language there. Finally, I put Singapore since I've always wanted to travel there. It's stunningly beautiful in all of the pictures that I've seen so to have the opportunity to work there and see the country for myself would be pretty amazing.

```
[ ]:
```