Real world data assignment

Directions: Use the following data sets to work on the questions:

EthnicFilePath = "data/race_and_ethnicity.csv"
EmploymentFilePath = "data/employment_by_occupations.csv"
GlobalDiversityPath = "data/global_diversity.csv"

1. Create a DataFrame of ethnic groups in New York, NY grouped by year and population. Sort data by year (ascending) and population (descending). Use data set: EthnicFilePath.
   HINT: Use pandas function read_csv to read a csv file.
2. Create a treemap with plotly express with the 1st question DataFrame.
   HINT: Use plotly express function treemap and pandas reset_index function (plotly express needs a DataFrame with no levels).
3. Create a histogram of employment by occupation. Use data set: EmploymentFilePath.
   HINT: Use plotly express function histogram.
4. Create a line chart representing the workforce in "Computer & Mathematical Occupations" from 2013 to 2019. Use data set: EmploymentFilePath.
   HINT: Filter the data by "Computer & Mathematical Occupations."
5. **BONUS:** Create a world map representing the birthplace of the foreign-born residents of New York, NY, through the years. Use data set: GlobalDiversityPath.
   HINT: Use Plotly Express function choropleth with the parameter animation_frame='year'.

List of functions you will need for this assignment:

Pandas (import pandas as pd):

- Read csv: df = pd.read_csv(FilePath)
- Group data: df_groupby = DataFrame.groupby(by=None, axis=0, level=None, as_index=True, sort=True, group_keys=True, squeeze=NoDefault.no_default, observed=False, dropna=True)
- Access group of row and columns: DataFrame.loc

Plotly Express (import plotly.express as px):

- Treemap charts: fig = px.treemap(DataFrame, path=[], values='', hover_data = [], color='')
- Histograms: px.histogram(DataFrame, x='', y='', color='')
- Line charts: px.line(DataFrame, x='', y='')
- Maps = px.choropleth(DataFrame, locations='', animation_frame='', color='')
- HTML versions of charts: fig.write_html(FilePath)

Documentation resources:

Python: https://docs.python.org/3/
Pandas: https://pandas.pydata.org/docs/
Plotly Express: https://plotly.com/python/plotly-express/