Output Description

The output will be a zipped file ('.zip' format) consisting of an Excel workbook (with the name '*inputfilename*_results.xlsx' with the following sheets. The sample output file attached here is for the May 2021 Basic Monthly CPS dataset.

General points:

- All metrics reported are not adjusted for seasonality and geography
- The logic of calculating the metrics is drawn from documents published by the U.S. Census Bureau
- Key references are as follows: https://www2.census.gov/programs-surveys/cps/datasets/2021/basic/2021_Basic_CPS_Public_Use_Record_Layout_plus_IO_Code_list.txt, https://www2.census.gov/programs-surveys/cps/methodology/PublicUseDocumentation_final.pdf

1. 'Households_ST': Provides the estimate of the aggregate number of households grouped by State.

    - Column 'Count' provides an estimate of aggregate number of households.
    - The logic of calculation is as follows. Each unique household in the dataset is multiplied by its respective weight 'HWHHWGT', then divided by 10000, and aggregated. This logic is based on the information provided here: https://www2.census.gov/programs-surveys/cps/methodology/PublicUseDocumentation_final.pdf.
    - Visualization of the 'Percentage of Total' is provided in the state-level U.S. Map.
    - For example, in the sample output file, the state 'AL' has an estimate of 2022644 households i.e., 1.55% of the total U.S. estimate.

2. 'Households_CSA': Provides the estimate of the aggregate number of households grouped by Consolidated Statistical Areas.

    - The same logic explained in the sheet 'Households_ST' applies here.
    - Visualization of the 'Percentage of Total' is provided in the bar chart.

3. 'Households_CBSA': Provides the estimate of the aggregate number of households grouped by Core-Based Statistical Areas.

    - The same logic explained in the sheet 'Households_ST' applies here.
    - Visualization of the 'Percentage of Total' is provided in the bar chart, grouped by U.S. Census Bureau Regions.
    - The mapping of CBSA with Regions is based on the information provided in two source: https://www2.census.gov/programs-surveys/metro-micro/geographies/reference-files/2020/delineation-files/list1_2020.xls and https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf.

4. 'Households_CO': Provides the estimate of the aggregate number of households grouped by Counties.

    - The same logic explained in the sheet 'Households_ST' applies here.

- Visualization of the 'Percentage of Total' is provided in the bar chart, grouped by U.S. Census Bureau Regions.
- The mapping of Counties with Regions is based on the information provided in two sources: https://www2.census.gov/programs-surveys/metro-micro/geographies/reference-files/2020/delineation-files/list1_2020.xls and https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf.

5. 'Household Members_ST': Provides the estimate of number of persons living in the households grouped by State.

    - The column 'Mean': Average of number of persons, living in the households, reported in the monthly CPS survey. This is captured using HRNUMHOU.
    - The column 'Weighted Mean': Estimate of average of number of persons living in the households i.e., average of the number of persons in each household multiplied by the respective household weight (i.e., HWHHWGT/10000).
    - The remaining columns 'Weighted Standard Deviation', 'Weighted Skewness', 'Weighted Kurtosis': Similar to the previous column, these columns are an estimate of the standard deviation, skewness, and kurtosis respectively i.e., all the 'weighted' columns are based on population estimates.
    - Visualization of the 'Weighted Mean' is provided in the state-level U.S. Map.

6. 'LaborForce_ST': Provides the estimate of Unemployment Rate, Employment-Population Ratio, Labor Force Participation Rate grouped by State. The logic of calculation of each column, explained below, is based on the information provided here: https://www2.census.gov/programs-surveys/cps/methodology/PublicUseDocumentation_final.pdf. For these employment metrics, only the civilian noninstitutional population (i.e., PRPERTYP = 2 and PRTAGE >= 16) is considered.

    - Column 'Unemployment Rate'

      Total number of unemployed people = if a person of age >= 16 and not member of U.S. Armed Forces is unemployed (i.e., PEMLR is 3 or 4), the statistical weight PWCMPWGT corresponding to that person is identified and aggregated.

      Total number of people in the labor force = if a person of age >= 16 and not member of U.S. Armed Forces is in the labor force (i.e., PEMLR is 1, 2, 3 or 4), the statistical weight PWCMPWGT corresponding to that person is identified and aggregated.

      Unemployment Rate = (Total number of unemployed people/Total number of people in the labor force)*100

    - Column 'Employment-Population Ratio'

      Total number of people employed = if a person of age >= 16 and not member of U.S. Armed Forces is employed (i.e., PEMLR is 1 or 2), the statistical weight PWCMPWGT corresponding to that person is identified and aggregated.

Civilian noninstitutional population = if a person is in the civilian noninstitutional population (i.e., PRPERTYP = 2 and age >= 16), the statistical weight PWCMPWGT corresponding to that person is identified and aggregated.

Employment-Population Ratio = (Total number of people employed/ Civilian noninstitutional population) * 100

- Column 'Labor Force Participation Rate'

  (Total number of people in the labor force/ Civilian noninstitutional population) * 100

- Column 'Number of Employed People': Total number of people employed (i.e., PEMLR is 1 or 2) in the civilian noninstitutional population.
- Column 'Number of People not in Labor Force': Total number of people out of labor force (i.e., PEMLR is 5, 6, or 7) in the civilian noninstitutional population.
- Column 'Number of Unemployed People': Total number of people unemployed (i.e., PEMLR is 3 or 4) in the civilian noninstitutional population.

  Note: Columns 'Number of Employed People', 'Number of People not in Labor Force', 'Number of Unemployed People' are not population estimates because they are not multiplied by the respective weights.

- Visualization of Unemployment Rate, Employment-Population Ratio, Labor Force Participation Rate is provided in separate state-level U.S. maps.

7. 'LaborForce_CSA', 'LaborForce_CBSA', 'LaborForce_CO': These sheets provide the estimate of Unemployment Rate, Employment-Population Ratio, Labor Force Participation Rate grouped by Consolidated Statistical Areas, Core-Based Statistical Areas, and Counties respectively.

   - The columns and their logic of calculation is the same as explained in the sheet 'LaborForce_ST'.
   - Visualization of Unemployment Rate, Employment-Population Ratio, Labor Force Participation Rate is provided in separate bar charts.
   - For Core-Based Statistical Areas and Counties, the visualization is provided in the bar chart, grouped by U.S. Census Bureau Regions.

8. 'LaborForce_Education', 'LaborForce_Sex', 'LaborForce_Race', 'LaborForce_Age':

   - The columns and their logic of calculation is the same as explained in the sheet 'LaborForce_ST'.
   - Visualization of Unemployment Rate, Employment-Population Ratio, Labor Force Participation Rate is provided in separate bar charts.
   - In the sheet 'LaborForce_Age', the age brackets were decided based on the documentation available here: https://www.census.gov/prod/2006pubs/tp-66.pdf.

9. 'FamilyIncome_ST': Provides the estimate of family income grouped by State.

- Column ‘Count’: Each household corresponding to their family income bracket is multiplied by their respective household weights (i.e., HWHHWGT/10000) and aggregated.
- Column ‘Percentage of State Total’: This provides the percentage of households with each family income bracket at a state-level. For example, in the sample output file, the State ‘AL’ has 3.7% households with a family income less than $5,000, 12.14% households with a family income between $75,000 to $99,999 and so on.
- Visualization of the maximum ‘Percentage of State Total’ within each State is provided in the state-level U.S. Map. For example, in the sample output file, consider the state AL: highest percentage of households (i.e., 12.14%) earn a family income between $75,000 to $99,999. In the state AK: highest percentage of households (i.e., 18.19%) earn a family income between $100,000 to $149,999. These metrics are presented in the state-level U.S. Map.

10. ‘Interview Status_ST’: Provides information on interview status (i.e. response rate and non-response rate) grouped by State.

    - Column ‘Interview Status’: This has four values – INTERVIEW, TYPE A NON-INTERVIEW, TYPE B NON-INTERVIEW, TYPE C NON-INTERVIEW.
    - Column ‘Count’: This is an aggregate of number of households in the survey grouped by their Interview Status.
    - Column ‘Percentage of State’: For example, in the sample output file, consider State ‘AL’. The value of ‘INTERVIEW’ for this state denotes that 73.34% of households participated in the monthly CPS survey i.e., response rate was 73.34%. Three types of non-response or non-interview (i.e., households refusing to participate in the survey) are quantified as Type A, Type B, and Type C.
    - Visualization of the ‘Percentage of State’ for ‘INTERVIEW’ value (i.e., response rate), within each State is provided in the state-level U.S. Map.

11. ‘Type A Non-Interview_ST’: Provides information on Type A non-response rate grouped by State.

    - Column ‘Type A Non-Interview Reason’: This has six values i.e., six reasons for Type A non-interview
    - Column ‘Count’: This is an aggregate of number of households in the survey grouped by the reasons for Type A non-interview.
    - Column ‘Percentage of State’: For example, in the sample output file, consider State ‘AL’. The value of ‘REFUSED (REF)’ for this state denotes that out of the households that did not participate in the monthly CPS survey due to ‘TYPE A Non-Interview’, 68.24% of them refused to participate.
    - Visualization of the ‘Percentage of State’ for ‘REFUSED (REF)’ value, within each State is provided in the state-level U.S. Map.

12. ‘Type B Non-Interview_ST’: Provides information on Type B non-response rate grouped by State.

    - Column ‘Type B Non-Interview Reason’: This has nine values i.e., nine reasons for Type B non-interview

- Column ‘Count’: This is an aggregate of number of households in the survey grouped by the reasons for Type B non-interview.
- Column ‘Percentage of State’: For example, in the sample output file, consider State ‘AL’. The value of ‘VACANT REGULAR’ for this state denotes that out of the households that did not participate in the monthly CPS survey due to ‘TYPE B Non-Interview’, 71.43% of them were found to be regularly vacant.
- Visualization of the ‘Percentage of State’ for ‘VACANT REGULAR’ value, within each State is provided in the state-level U.S. Map.

13. ‘Type C Non-Interview_ST’: Provides information on Type C non-response rate grouped by State.

    - Column ‘Type C Non-Interview Reason’: This has ten values i.e., ten reasons for Type C non-interview
    - Column ‘Count’: This is an aggregate of number of households in the survey grouped by the reasons for Type C non-interview.
    - Column ‘Percentage of State’: For example, in the sample output file, consider State ‘AL’. The value of ‘DEMOLISHED’ for this state denotes that out of the households that did not participate in the monthly CPS survey due to ‘TYPE C Non-Interview’, 28.57% of them were found to be demolished.
    - Visualization of the ‘Percentage of State’ for ‘DEMOLISHED’ value, within each State is provided in the state-level U.S. Map.

14. ‘Hourly Earnings_ST’: Provides estimate of hourly earnings grouped by State.

    - The column ‘Mean’: Average of hourly earnings of persons in the monthly CPS survey. This is captured using (PTERNHLY/100).
    - The column ‘Weighted Mean’: Estimate of hourly earnings of persons i.e., average of the hourly earnings of each person multiplied by the respective second stage weight (i.e., PWSSWGT/10000).
    - The remaining columns ‘Weighted Standard Deviation’, ‘Weighted Skewness’, ‘Weighted Kurtosis’: Similar to the previous column, these columns are an estimate of the standard deviation, skewness, and kurtosis respectively i.e., all the ‘weighted’ columns are based on population estimates.
    - Visualization of the ‘Weighted Mean’ is provided in the state-level U.S. Map.

15. ‘Hourly Earnings_CSA’, ‘Hourly Earnings_CBSA’, ‘Hourly Earnings_CO’: These sheets provide the estimate of hourly earnings grouped by Consolidated Statistical Areas, Core-Based Statistical Areas, and Counties respectively.

    - The columns and their logic of calculation is the same as explained in the sheet ‘Hourly Earnings_ST’.
    - Visualization of ‘Weighted Mean’ is provided in separate bar charts.
    - For Core-Based Statistical Areas and Counties, the visualization is provided in the bar chart, grouped by U.S. Census Bureau Regions.

16. ‘Hourly Earnings_Education’, ‘Hourly Earnings_Sex’, ‘Hourly Earnings_Race’, ‘Hourly Earnings_Age’: These sheets provide the estimate of hourly earnings grouped by education level, sex, race, and age respectively.

    - The columns and their logic of calculation is the same as explained in the sheet ‘Hourly Earnings_ST’.
    - Visualization of ‘Weighted Mean’ is provided in separate bar charts.
    - In the sheet ‘Hourly Earnings_Age’, the age brackets were decided based on the documentation available here: https://www.census.gov/prod/2006pubs/tp-66.pdf.

17. ‘Weekly Earnings_ST’: Provides estimate of weekly earnings grouped by State.

    - The column ‘Mean’: Average of weekly earnings of persons in the monthly CPS survey. This is captured using (PTERNWA/100).
    - The column ‘Weighted Mean’: Estimate of weekly earnings of persons i.e., average of the hourly earnings of each person multiplied by the respective second stage weight (i.e., PWSSWGT/10000).
    - The remaining columns ‘Weighted Standard Deviation’, ‘Weighted Skewness’, ‘Weighted Kurtosis’: Similar to the previous column, these columns are an estimate of the standard deviation, skewness, and kurtosis respectively i.e., all the ‘weighted’ columns are based on population estimates.
    - Visualization of the ‘Weighted Mean’ is provided in the state-level U.S. Map.

18. ‘Weekly Earnings_CSA’, ‘Weekly Earnings_CBSA’, ‘Weekly Earnings_CO’: These sheets provide the estimate of weekly earnings grouped by Consolidated Statistical Areas, Core-Based Statistical Areas, and Counties respectively.

    - The columns and their logic of calculation is the same as explained in the sheet ‘Weekly Earnings_ST’.
    - Visualization of ‘Weighted Mean’ is provided in separate bar charts.
    - For Core-Based Statistical Areas and Counties, the visualization is provided in the bar chart, grouped by U.S. Census Bureau Regions.

19. ‘Weekly Earnings_Education’, ‘Weekly Earnings_Sex’, ‘Weekly Earnings_Race’, ‘Weekly Earnings_Age’: These sheets provide the estimate of weekly earnings grouped by education level, sex, race, and age respectively.

    - The columns and their logic of calculation is the same as explained in the sheet ‘Weekly Earnings_ST’.
    - Visualization of ‘Weighted Mean’ is provided in separate bar charts.
    - In the sheet ‘Weekly Earnings_Age’, the age brackets were decided based on the documentation available here: https://www.census.gov/prod/2006pubs/tp-66.pdf.

20. ‘Major Industry_ST’, ‘Major Industry_CSA’: These sheets provide information on employment (primary job) in major industries grouped by State and Consolidated Statistical Areas. This is captured using PRMJIND1.

    - Column ‘Major Industry’: There are 14 types of major industries.

- Column ‘Count’: Estimate of number of people employed in each type of major industries i.e., aggregate of the number of persons multiplied by the respective second stage weight (i.e., PWSSWGT/10000).
- Column ‘Percentage of State’: For example, in the sample output file, consider the State AL. The numbers denote that AL has 472819 people employed in Educational and Health Services. This comprises 21.47% of the employment in the state AL.
- Visualization of the ‘Major Industry’ with the highest ‘Percentage of State’ is provided in the state-level U.S. Map.

21. ‘Intermediate Industry_ST’, ‘Intermediate Industry_CSA’: These sheets provide information on employment (primary job) in intermediate industries grouped by State and Consolidated Statistical Areas. This is captured using PRIMIND1.

    - Column ‘Intermediate Industry’: There are 22 types of intermediate industries.
    - Column ‘Count’: Estimate of number of people employed in each type of intermediate industries i.e., aggregate of the number of persons multiplied by the respective second stage weight (i.e., PWSSWGT/10000).
    - Column ‘Percentage of State’: For example, in the sample output file, consider the State AL. The numbers denote that AL has 278703 people employed in Health Care and Social Assistance Services. This comprises 12.66% of the employment in the state AL.
    - Visualization of the ‘Intermediate Industry’ with the highest ‘Percentage of State’ is provided in the state-level U.S. Map.

22. ‘Major Occupation Categories_ST’, ‘Major Occupation Categories_CSA’: These sheets provide information on employment in major occupation categories grouped by State and Consolidated Statistical Areas. This is captured using PRMJOCGR.

    - Column ‘Major Occupation Category’: There are 7 types of major occupation categories.
    - Column ‘Count’: Estimate of number of people employed in each type of major occupation categories i.e., aggregate of the number of persons multiplied by the respective second stage weight (i.e., PWSSWGT/10000).
    - Column ‘Percentage of State’: For example, in the sample output file, consider the State AL. The numbers denote that AL has 767843 people employed in Management, Professional, and Related Occupations. This comprises 34.87% of the employment in the state AL.
    - Visualization of the ‘Major Occupation Category’ with the highest ‘Percentage of State’ is provided in the state-level U.S. Map.

23. ‘Major Occupation_ST’, ‘Major Occupation_CSA’: These sheets provide information on employment (primary job) in major occupations grouped by State and Consolidated Statistical Areas. This is captured using PRMJOCC1.

    - Column ‘Major Occupation’: There are 11 types of major occupations.
    - Column ‘Count’: Estimate of number of people employed in each type of major occupations i.e., aggregate of the number of persons multiplied by the respective second stage weight (i.e., PWSSWGT/10000).

- Column ‘Percentage of State’: For example, in the sample output file, consider the State AL. The numbers denote that AL has 457241 people employed in Professional and Related Occupations. This comprises 20.76% of the employment in the state AL.
- Visualization of the ‘Major Occupation’ with the highest ‘Percentage of State’ is provided in the state-level U.S. Map.

24. ‘Country Totals’: This sheet provides a snapshot of key metrics at the national level. Interpretation of these key metrics and their logic of calculation are consistent with the explanations provided in the sheets above.