SPP Empowering Data

Exploratory Data Analysis Template

Name:

Part 3 Instructions

1. Open the LivWell135 Column descriptions document (There is also a LivWell175 document with the column descriptions for that dataset on HQ.)
2. Search through the descriptions for either columns that you are interested in knowing more about or that you think might be related somehow.

   Remember that there are no “right” or “wrong” choices of columns - this is about finding what you are interested in or where you think relationships might exist.

3. When you find a column, copy **both** the column name **and** the column description to the one of the Column Information tables of your copy of this document. You are only filling in the *column name* and *description* sections of each table. Be sure to copy the column name *exactly* as it appears - we’ll be using it in our code later and need an exact match.
4. Aim to find 3 to 5 columns. If you are interested in more, you can add more spaces in your document.

Remember, for Part 3, you are only filling in the *Column name* and *Column description* sections of each table.

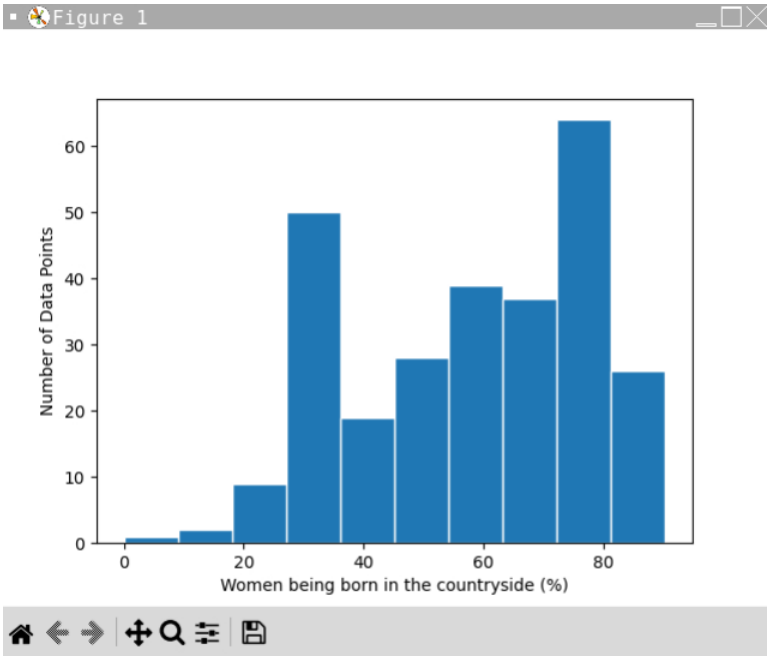
Part 4 Instructions

For each of the columns you chose to explore, do the following:

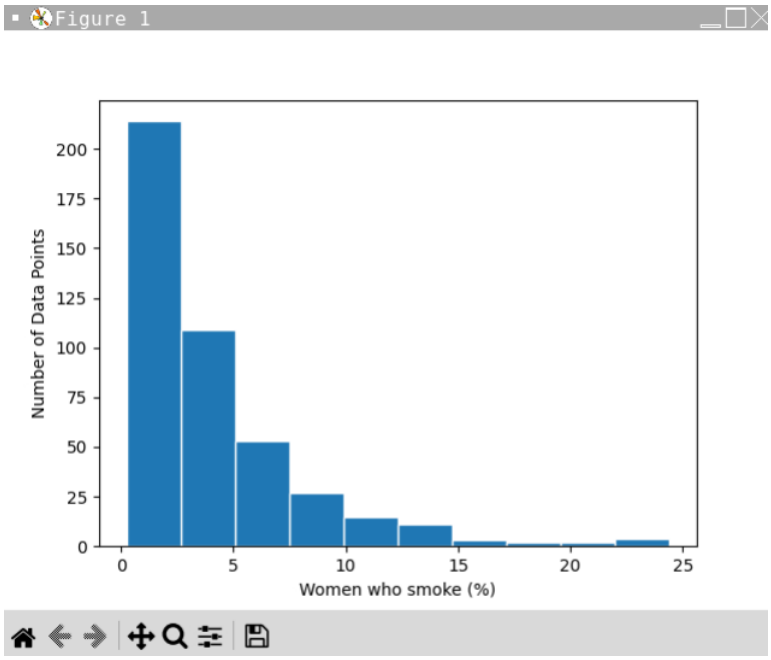
1. Modify the x parameter in the `hist` method to use the column name you noted in your table.
2. Modify the `xlabel` method to use the column description string you noted in your table.
3. Rerun the replit to generate a histogram of the data in that column.
4. Add a screenshot of the histogram to your Column Information table. Here’s a video with instructions on how to do it on Chromebooks, PCs and Macs.
5. Also record the maximum, minimum and most common value information to the table. You don’t need exact values - approximate values are good enough.
6. Repeat this for all the columns you chose.

SPP Empowering Data

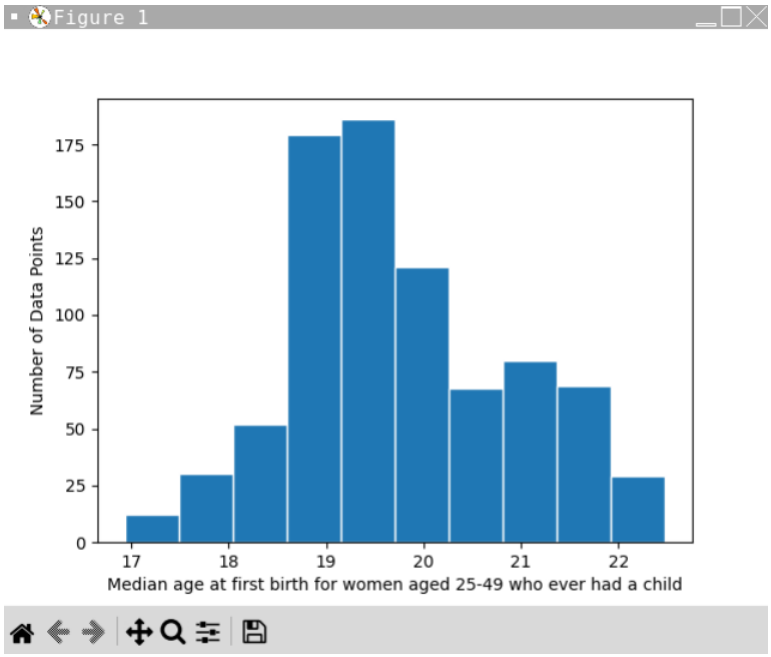
Exploratory Data Analysis Template

Chosen column 1	
Column name (label)	DM_born_rural_p
Column description	Women being born in the countryside (%)
Histogram	
Maximum value in column	90
Minimum value in column	1
Most common value in column (tallest histogram bar)	75
Mean (or median) for column	

Chosen column 2	
Column name (label)	HL_smoke_p
Column description	Women who smoke (%)

Histogram	
Maximum value in column	24
Minimum value in column	0
Most common value in column (tallest histogram bar)	2
Mean (or median) for column	

Chosen column 3	
Column name (label)	RH_age_first_birth_mean
Column description	Median age at first birth for women aged 25-49 who ever had a child

Histogram	
Maximum value in column	22
Minimum value in column	17
Most common value in column (tallest histogram bar)	19
Mean (or median) for column	