Cleaning the Data

Hayaku

03/05/2021

Research Question 1

Questions changed to:

How has the popularity of each type of alcohol changed over the years for different countries and have the local favorites of earlier decades been overtaken by other types of alcohol in terms of consumption?

Researcher: Yalong Liu

Data of all recorded timelines

We need to connect the timelines of different data sets for observation. The raw data set of consumption from 1960 to 2018 is cleaned.

Data by gender

This data is used to view the total alcohol consumption per capita of different genders in various countries in recent years, so it is necessary to meet the requirements of screening by gender and timeline at the same time.

Combine this data set with the data set in the first part to explore whether alcoholic beverages are substituted.

Research Question 2

Does alcohol take more lives in traffic crashes or illnesses and which countries suffer more from alcohol-related cases?

Researcher: Yin Shan Ho

Alcohol-attributed deaths & Number of death per 100,000 population

Variables are renamed and "no data" values are filtered out. Pivot longer is used to combine the separated variables on gender into one for future calculations.

Research Question 3

What are the recurring patterns of policy responses across different countries in the world and how has it impacted the consumption of alcohol?

Researcher: Smriti Vinayak Bhat

Data Pertaining to Policy Responses

This filters out columns for the regions with no data. The data from all countries is then combined by year and action area using a full join. Finally all the countries' policy pattern together are bound together using rbind.

Data Pertaining to Trend of Consumption of Alcohol

The columns are renamed for convenience.The variables X2 and 2018 are dropped as considered not necessary for the report. The first row is dropped as it is a repeat of the heading. Then countries are filtered out for whom policy data is not present.

Research Question 4

Do drinking age restrictions and legal blood alcohol concentration (BAC) limits for young drivers have any effect on adolescents' drinking among countries?

Researcher: Kexin Xu

Cleaning the Data

First row is dropped for containing redundant values and variables are renamed appropriately according to value for all data. This is then pivoted into long form and unnested or seperated to get correct variables.

Research Question 5

What are the trends in youth drinking issues over the years?

Researcher: Kexin Xu

Cleaning the Data

The trend data is cleaned by removing redundant rows and appropriate renaming. Then it pivoted longer into where all the variables except country and year are placed into gender variable with values in percentage.