

# Coffee and code

## Introduction

Coffee is a very common and popular drink because the caffeine in coffee can help people relieve fatigue. For those who need to sit in the office for a long time, drinking a few cups of coffee is normal. Especially in the IT industry, IT developers need to program for a long time or even all night, so coffee can play a role in this situation.

This is a data set "Coffee and code dataset" from Kaggle with the theme of coffee and coding. This data set collects samples of 100 people, who are coffee lovers and code developers living in Lebanon. This data reveals information about them, such as coding hours and the average number of coffee drinks per day.

I will use this data for visualization in this article and try to find the direct relationship between coffee and coding.

## Data Overview

This data contains the following information.

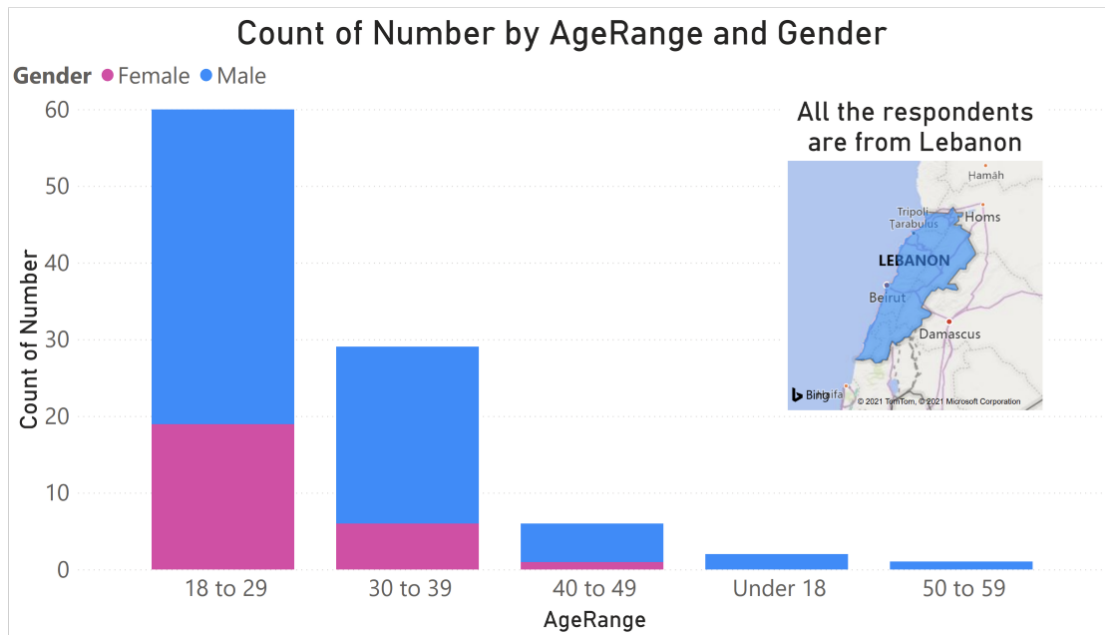
"CodingHours", "CoffeeCupsPerDay", "CoffeeTime", "CodingWithoutCoffee", "CoffeeType", "CoffeeSolveBugs", "Gender", "Country", "AgeRange".

The "Gender", "Country", "AgeRange" are the general information of the respondents.

The "CodingHours", "CoffeeSolveBugs" be able to reflect the work performance of the investigated. If coffee can make them more energetic, then "CodingHours" will reflect this. Similarly, if coffee can make people's coding work better, "CoffeeSolveBugs" will reflect this. The "CoffeeType" "CoffeeCupsPerDay" "CoffeeTime" "CodingWithoutCoffee" shows how they drink coffee.

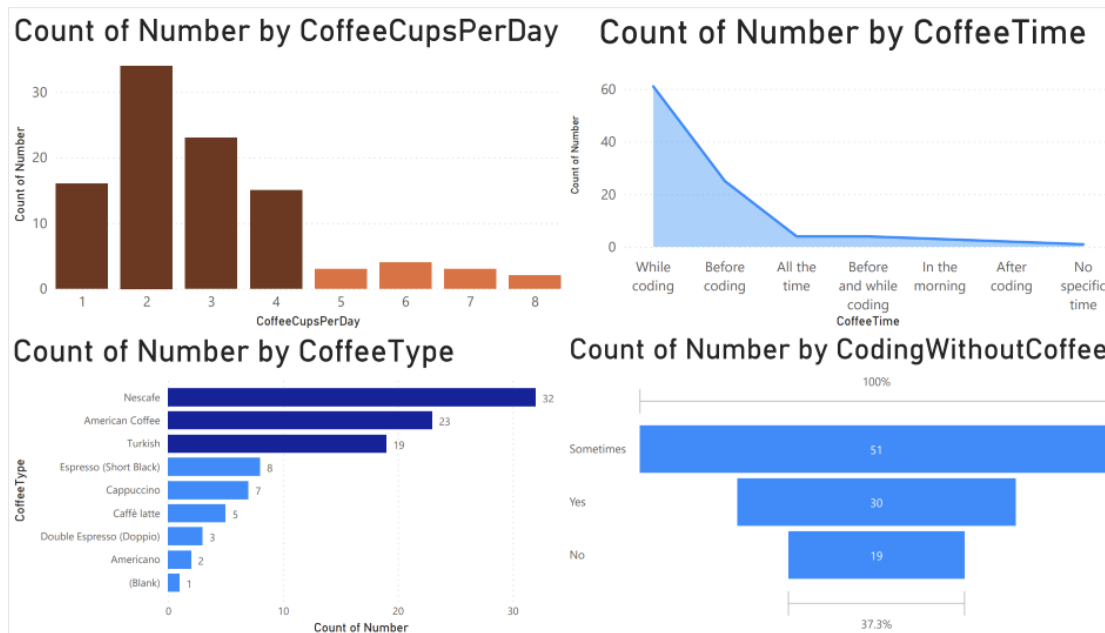
## Data Analysis

As shown in "Picture from Power Bi 1", The distribution of gender proportion in the sample is unbalanced, and the male sample(72) is significantly more than the female sample(26). The age distribution in the sample is in line with the expectation. Most of the coding developers are between 18-39 years old, and the developers aged 18-29 account for nearly 40% of the total sample. However, there are too few female samples in the three intervals of "40 to 49", "under 18" and "50 to 59".



Picture from Power Bi 1

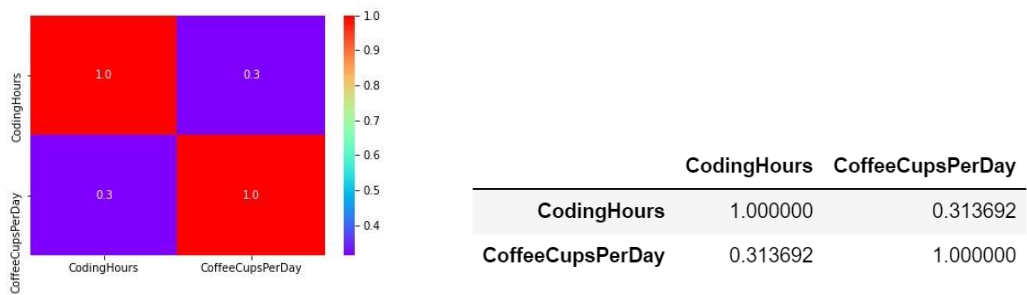
“Picture from power Bi 2” shows us the relevant information about the respondents drinking coffee. We can know that most people drink coffee during “while coding” and “before coding”, which may indicate that they really need coffee during coding, but the results of “CodingWithoutCoffee” also show that coffee is not necessary for coding, About 30 percent said they could code without coffee. Most people drink two or more cups of coffee a day, while more than five cups a day are rare. The “CoffeeType” also reflect the respondents' preference for three kinds of coffee.



Picture from Power BI 2

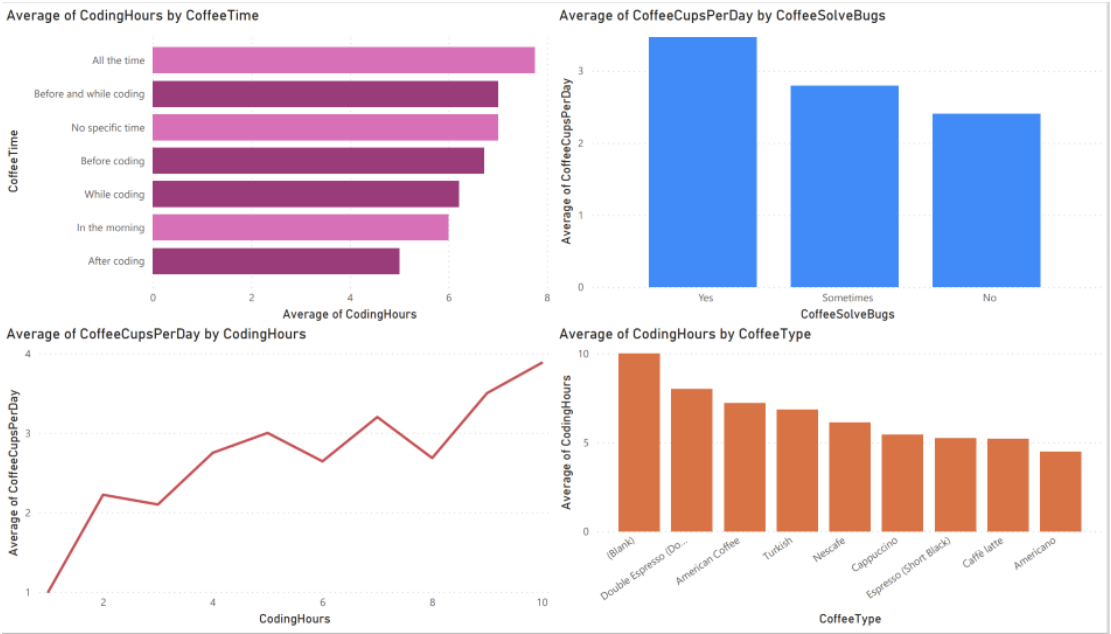
We can use Pearson correlation coefficient to find the relationship between these coffee variables and coding. As the “Picture from Python 1” show that There is a linear positive correlation between daily coffee consumption and coding hours. As the respondents spend

more time coding every day, they need to drink more coffee.



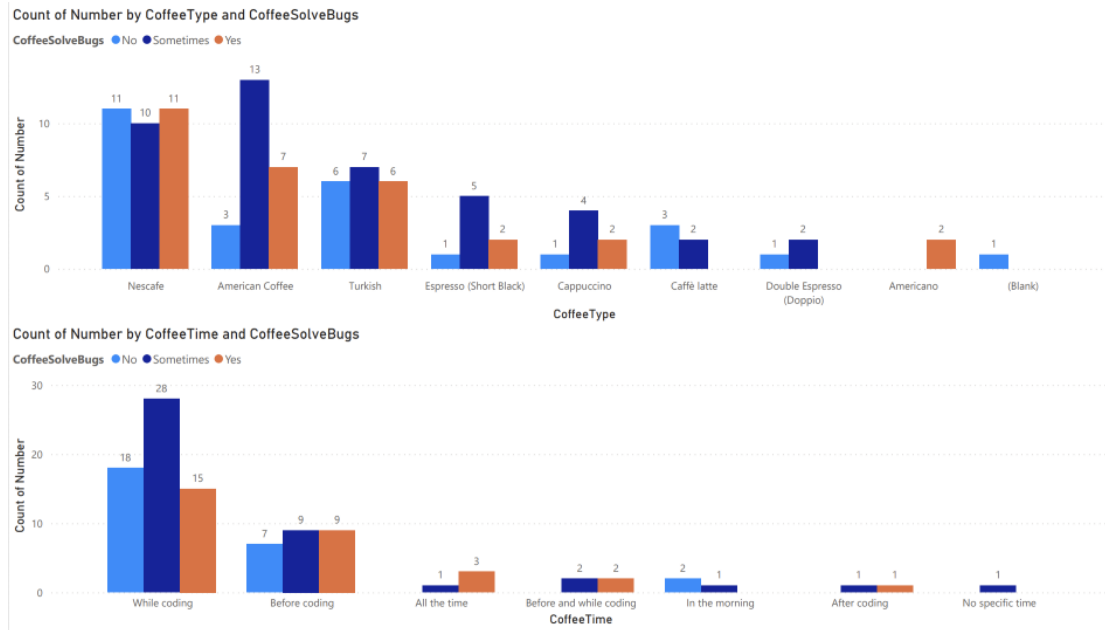
Picture from Python 1

The relationship between coffee and coding can be clearly drawn from the “picture from power Bi 3”. The graph “average of coffee cups per day by coding hours” shows the relationship between daily coffee intake and coding hours. This linear graph confirms the above positive correlation conclusion obtained through Pearson coefficient. We can know from the graph “average of coffee cups per day by coffee “SolveBugs”, People who can solve bugs better also tend to drink more coffee. In the first sub figure in the upper left corner, we observe four lines related to coding: “before and while coding”, “before coding”, “while coding”, “after coding”. It seems that the earlier you drink coffee during coding, the longer you will continue to work. Coffee types did not show a significant impact on “CodingHours” (The “Blank” means unknow).



Picture from Power Bi 3

Figure "picture from power Bi 4" tries to find the relationship between "CoffeeType", "CoffeeTime" and "CoffeeSolveBugs". Only American coffee shows that "yes" far exceeds "no", which means that solvingbugs are more capable among people who may drink American coffee. For other attributes, either the number of samples is too small, or "yes" is not much different from "no"



Picture from Power BI 4

## Summary

Some information can be learned from this data set. Increasing coffee intake can significantly improve people's coding time and slightly improve people's ability to solve bugs. Drinking coffee before and during coding can also increase people's coding hours. However, different brands of coffee do not reflect the impact of significant differences on people's coding.

## Appendix

The link of original dataset: [Coffee and code dataset | Kaggle](https://www.kaggle.com/shrutikunapuli/coffee-and-code-dataset)

Or (<https://www.kaggle.com/shrutikunapuli/coffee-and-code-dataset>)