

The Impacts of Marriage and Age in Hairdresser Salon Tipping Amounts

FREDERIK VINCENT PRIMDAHL MIKAEL STEENBERG PASOVSKI
ANDREAS PETER BRODERSEN ANDREAS WEDE GUSTAVSEN
Copenhagen School of Design and Technology

March 15, 2022

Abstract

In this study the data gathered from customers visiting a hairdresser salon is used to correlate tipping amount with a customer's civil status, age, culture, and the time of day of the visit, to assess which combination of measures affect tipping amount the most. It is proposed that the greatest variance in tipping amount is based on age and civil status, and that these correlate positively with tipping amount, with a married civil status and young age resulting in higher tips. In addition, this study describes the ethical dilemmas associated with datasets concerning culture. Through analyzation of the dataset, tipping amount is concluded to vary primarily based upon age, with civil status, culture and age as secondary contributing factors.

I. INTRODUCTION

Tipping is in many cultures a central aspect of the exchange of services, yet one often relying on implicit social norms which can vary greatly depending on factors like generational differences and culture. In this study we take a look at the influencing factors in the amount of tips received by the hairdresser in our dataset. This was done to examine tipping etiquette between different groups of people.

II. ETHICS

It is important when analysing the data from this study to keep ethics in mind. In this study we deal with data categorized by culture, and therefore a certain bias must be taken into consideration. Maintaining privacy and respecting human rights are just some of the ethical challenges that arise. Keeping ethics into consideration will ensure trust, and accountability to AI that could use the data.

III. METHODS

In our analysis we applied the data-driven data science paradigm as to find new insights from the provided haircut tip data set. As compared to hypothesis-driven data science where a hypothesis/problem is formulated and then data is collected or found to help answer the hypothesis.

This study makes use of the multi-methodological research framework put forth by Nunamaker, and follows four complementary research strategies to alleviate the complexity involved in different research activities. The strategies involve observation, for achieving an understanding of the domain, and to develop hypotheses based on this understanding, theory building in which we examined how to handle the questions from our observation, experimentation, where hypotheses are tested, and finally systems development, for designing and developing the final prototype showcasing the results of the research.

As opposed to a more linear scientific method, the multi-methodological approach supports repeating steps in order to continu-

ously get feedback from previous research activities and gain a more complete understanding of the research area.

For the analyzation of the haircut tipping amount dataset used in this research we make use of inferential statistics, to study a sample group and make generalizations about the wider population. As an extension of this, correlation analysis is also performed to determine how the different variables affect the tipping amount.

In the analysis of the haircut tip data set multiple data types are used, these data types restrict what kind of statistical tests can be applied on the data. Two overall categories are used: Non-parametric(categorical data, do not represent numerical data) and Parametric data. The Non-Parametric consists of Nominal, which contains labeled data and Ordinal data, which contains labeled data but in a specific order. Parametric consists of Interval, which is evenly spaced numerical data, where evenly spaced means the delta between points are the same. Ratio has the same properties as Interval but with a true zero representation, which means there is an interpretation connected to what true zero means.

In the haircut tip data set we can see examples of all 4 data types mentioned. An example of nominal can be seen in the "Married" column which contains the labels "M/NM". Ordinal can be seen in the "Day" column, labeled data in a specific order, monday to sunday. Interval can be seen in "Time" column and last Ratio can be seen in "Tip amounts" where true zero is interpreted as the absence of tipping.

IV. ANALYSIS

Data cleanup of the initial dataset in this study was needed. First all the undefined columns were removed. Secondly it was needed to format poorly formatted day labels, as these were inconsistent. As an example some were labelled "Tues" and others "Tue". Converting time to military time was also needed, since the data included a mixed set of 24-hour time and 12-hour time. This ended up in "5" being

rewritten as "17". Lastly we made time and age into integers instead of floats.

To be able to explore and analyse the data, we've chosen to illustrate it using a mix of line and bar diagrams. Figure 1 is the most promising with its clear separation in married and unmarried tip amounts. While figure 1 has been main focus of the study, figure 2 and 3 got shows outliers and are worth exploring.

V. FINDINGS

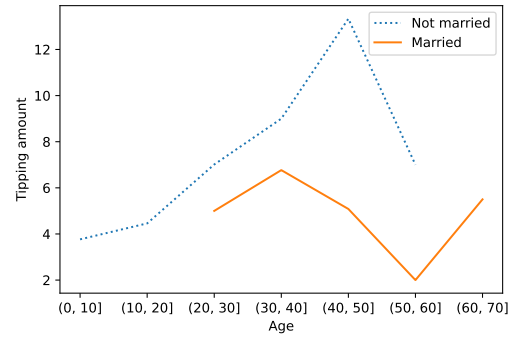


Figure 1: Mean tip amount split by marital status binned by age.

Figure 1 clearly indicates a correlation between marriage, age and tipping amount. At no age group do we see the mean of the married customers dips under the mean of the unmarried.

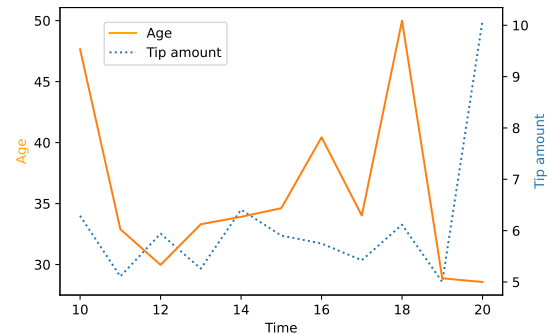


Figure 2: Mean tip amount split by marital status binned by time of day.

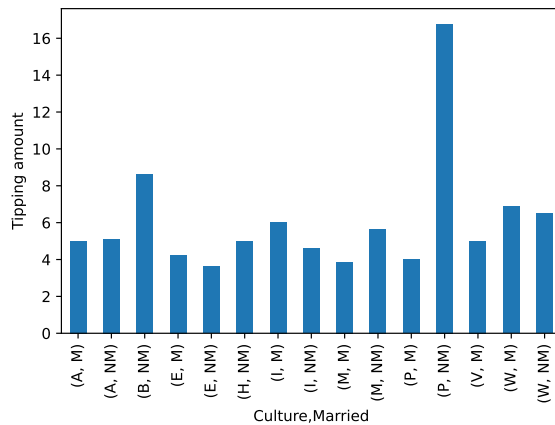


Figure 3: Mean tip amount split by marital status binned by culture.

VI. CONCLUSION

In the study of the haircut tip data set we have focused on whether tipping amount correlates with a customer's civil status, age, culture and the time of day. Through the study we found that there are patterns that correlates between marriage, age, and tipping amount. This can be seen in Figure1, here not married in the ages 40-50 tip significantly more than married in the same age group.

REFERENCES

- [Jay F. Nunamaker, Minder Chen and Titus D.M. Purdin.]
Systems Development in Information Systems Research". In: Journal manangement Information Systems (1990), pp. 90-96
- [Lecture 6.] URL <https://kea-teachers.gitlab.io/kea-aai-2022-1/lecture06.html>