

Article Title

FREDERIK VINCENT PRIMDAHL MIKAEL STEENBERG PASOVSKI

ANDREAS PETER BRODERSEN ANDREAS WEDE GUSTAVSEN

Copenhagen School of Design and Technology

March 14, 2022

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

I. 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 greater 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.

Lorem ipsum dolor sit amet, consectetur adipiscing elit.

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 con-

sideration. 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 continuously 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.

Maecenas sed ultricies felis. Sed imperdiet dictum arcu a egestas.

- Donec dolor arcu, rutrum id molestie in, viverra sed diam
- Curabitur feugiat
- turpis sed auctor facilisis
- arcu eros accumsan lorem, at posuere mi diam sit amet tortor
- Fusce fermentum, mi sit amet euismod rutrum
- sem lorem molestie diam, iaculis aliquet sapien tortor non nisi
- Pellentesque bibendum pretium aliquet

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Text requiring further explanation¹.

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

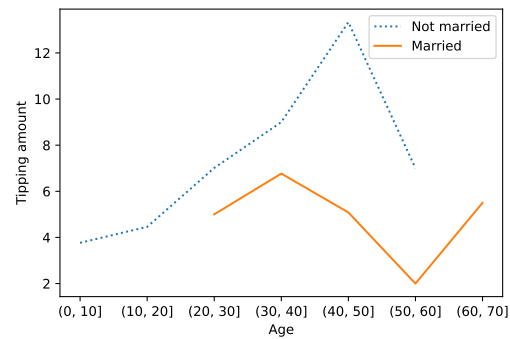


Figure 1: Mean tip amount split by marital status binned by age.

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.

¹Example footnote

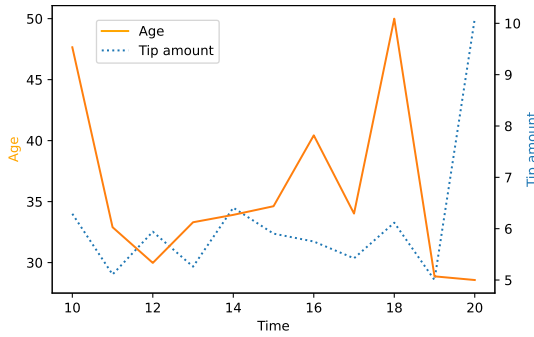


Figure 2: Mean tip amount split by marital status binned by time of day.

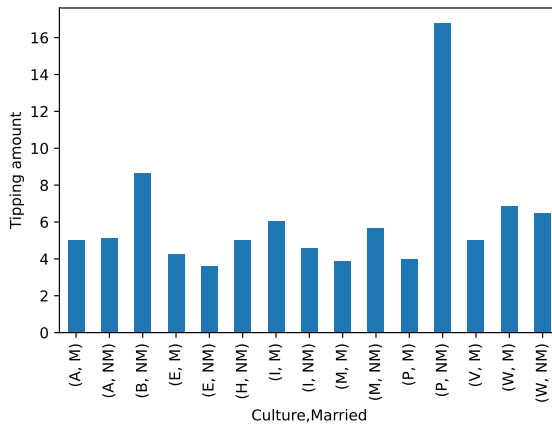


Figure 3: Mean tip amount split by marital status binned by culture.

VI. DISCUSSION

i. Subsection One

A statement requiring citation [Figueredo and Wolf, 2009]. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla

tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

ii. Subsection Two

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

REFERENCES

[Figueredo and Wolf, 2009] Figueredo, A. J. and Wolf, P. S. A. (2009). Assortative pairing and life history strategy - a cross-cultural study. *Human Nature*, 20:317–330.