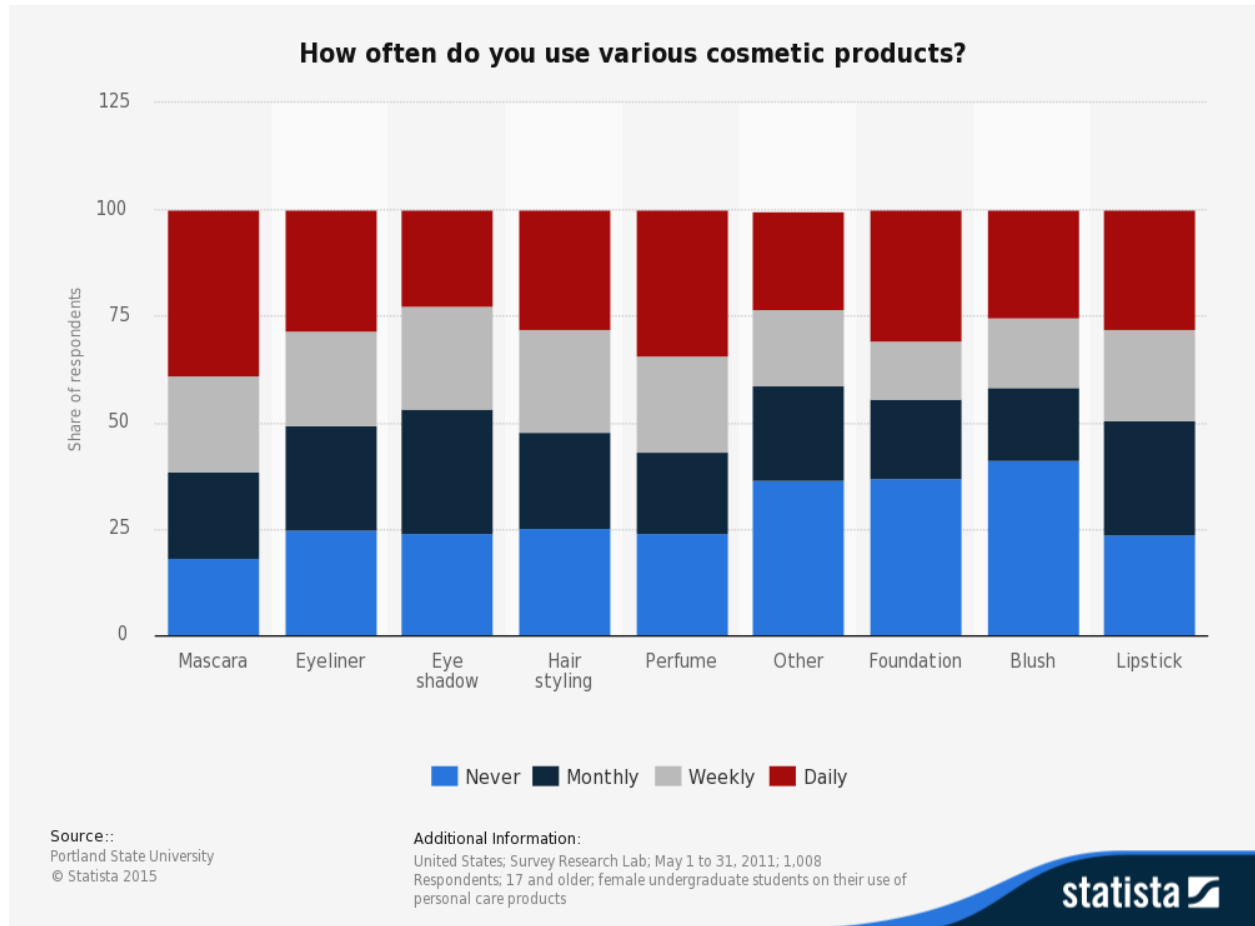


Comparing Expenses on Cosmetics By U.S. and Chinese College Students



(source: Statista, <http://www.statista.com/statistics/259345/frequency-of-use-of-cosmetic-products-by-college-students-in-the-us/>)

I recently studied a survey on the frequency of use of cosmetic products by college students in U.S. As shown in the grouped bar chart above, on average over 2/3 of American college students wear some kind of make-up on a monthly, weekly, or daily basis. As a college student who lived in China for the first half of her life, I am interested in studying the difference between American and Chinese students' semi-annual expenses on cosmetic products.

Research Question

I'm interested in the difference in mean expenses on cosmetics between American and Chinese college students. In essence, I want to have a glimpse of different, if so, consumer behaviors between two groups of students with disparate cultural background, and analyzed possible reasons behind the respective consumer behaviors.

Population of Interest

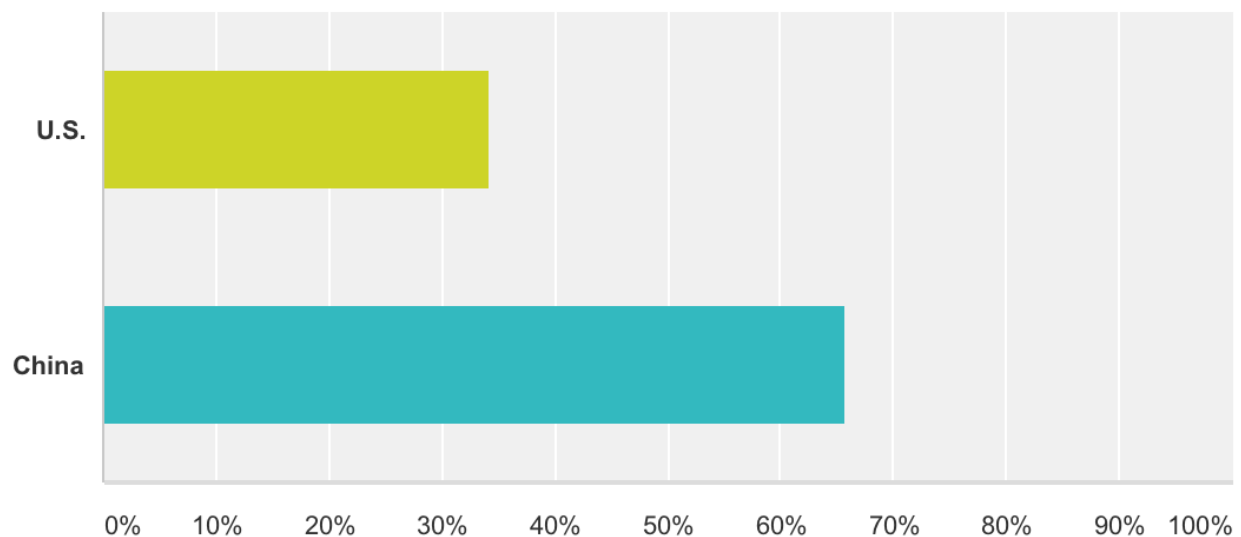
My population of interest is American and Chinese college students. I want to make a very important clarification about the definition of “American and Chinese college students”. “American college students” means students who grow up in U.S and now are college students, and similarly “Chinese students” means students who grow up in China and now are college students.

Data Collection

I started collecting data a week ago through SurveyMonkey and social media Facebook and WeChat (Chinese mainstream social media platform). I have a sample size of 70 college students with 24 of them growing up in the U.S. and 46 of them in China.

Where do you grow up?

Answered: 70 Skipped: 3



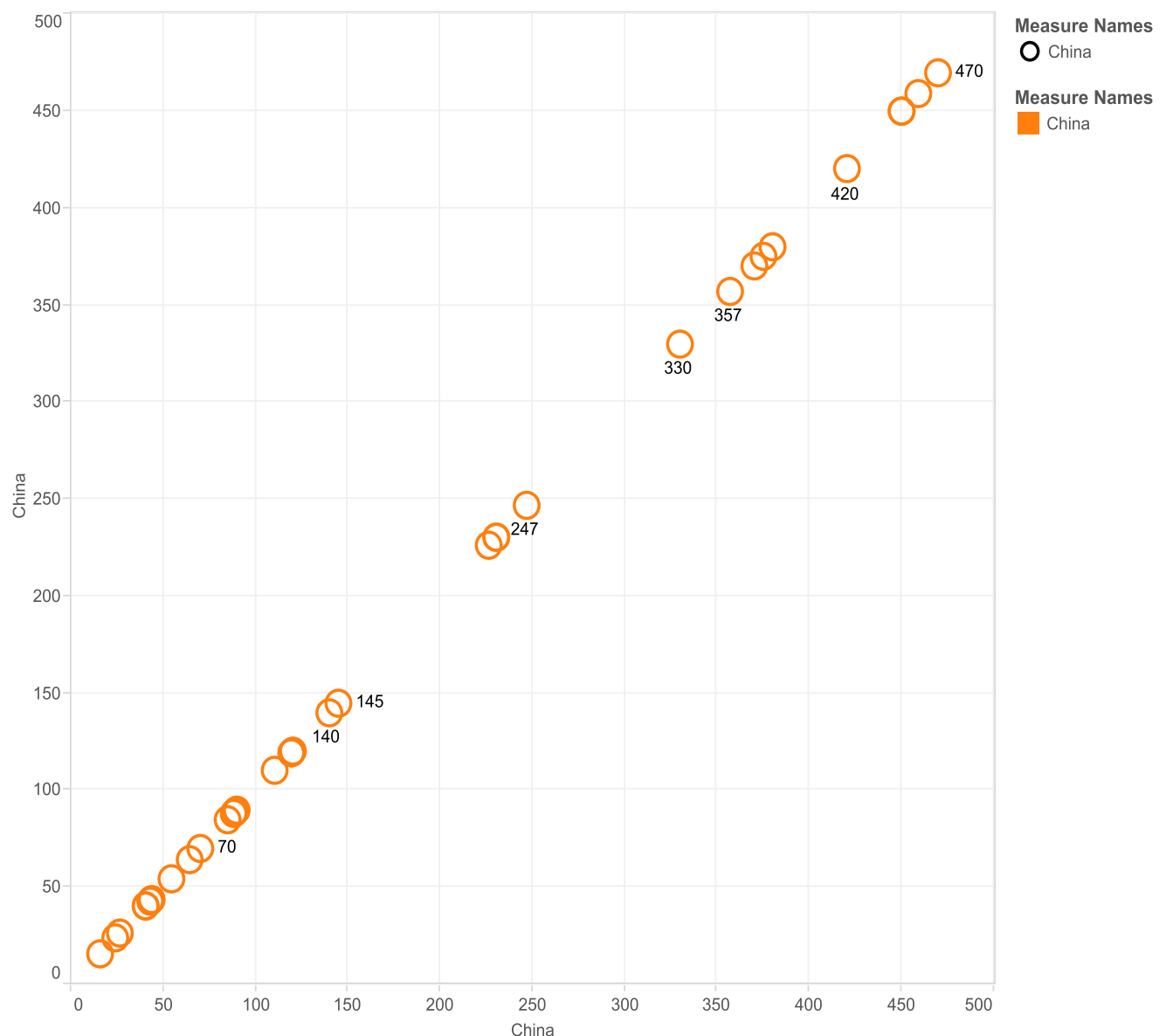
13 of the 70 respondents gave null information on their expenses, so I ultimately have an effective sample size of 57 respondents with 22 American and 35 Chinese. The following are data visualizations in Tableau.

Chinese college students' expenses on cosmetics semiannually: 5 Number Summary

Max: 470; Min: 16; Mean: 191.76476;

Median: 119.5; Q1: 65.5; Q3: 350.25

Sheet 1



China vs. China. Color shows details about China. Shape shows details about China. The marks are labeled by China.

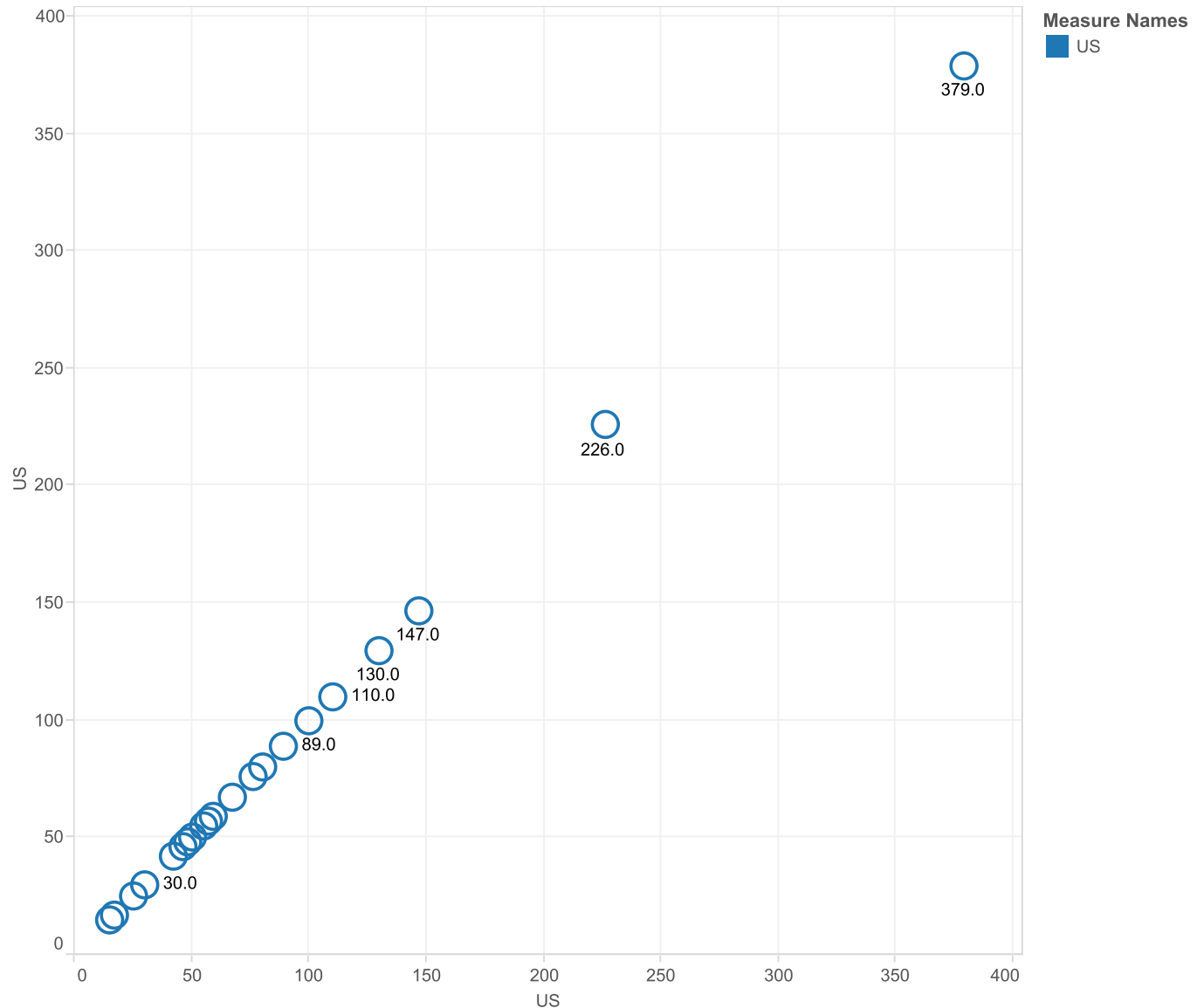
American college students' expenses on cosmetics semiannually:

5 Number Summary

Max: 379; Min: 17; Mean: 106.8;

Median: 63; Q1: 47.5; Q3: 115

Sheet 1



US vs. US. Color shows details about US. The marks are labeled by US. Details are shown for US.

Statistical Method

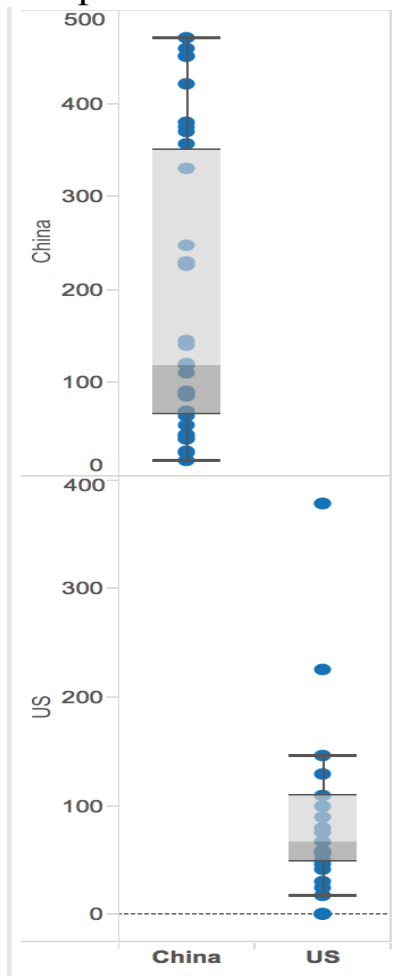
I want to know whether Chinese college students are likely to spend differently on cosmetics than American college students. I have dollars spent buying cosmetics in half a year for 57 randomly selected college students, 22 growing up in U.S. and 35 in China. I wonder whether the difference between mean expenses is zero.

Hypothesis

Null: The difference in mean expenses on cosmetics by American and Chinese college students is zero.

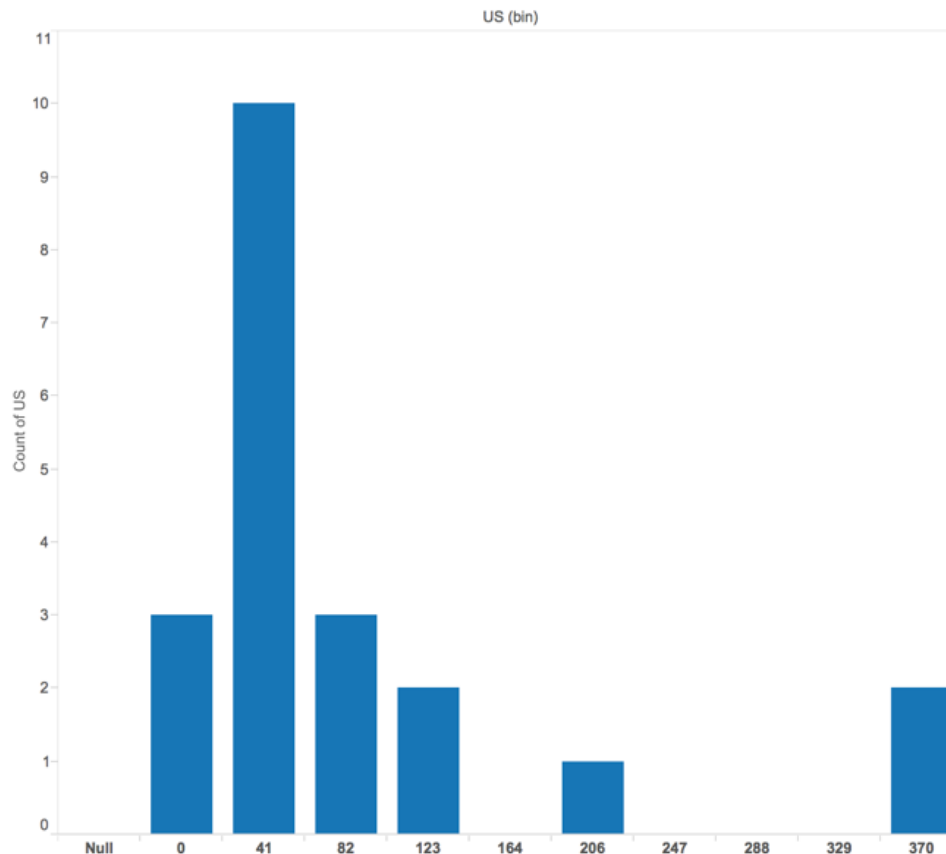
Alternative: The difference in mean expenses is not zero.

Boxplot

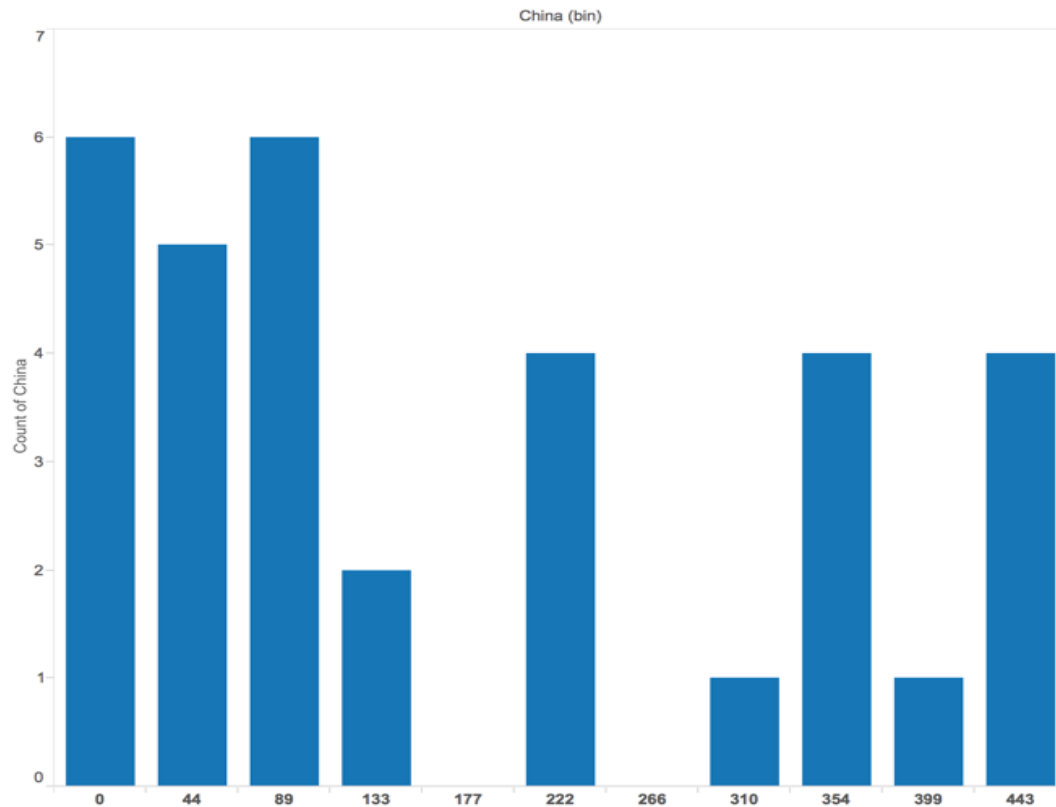


Checking Assumptions and Conditions

- Randomization Condition: The experiment was randomized. Subjects were assigned to treatment groups at random.
- Independence Assumption: This is an experiment, so there is no need for the subjects to be randomly selected from any particular population. All we need to check is whether they were assigned randomly to treatment groups.
- Independent Groups Assumption: Randomizing the experiment gives independent groups.
- Nearly Normal Condition: Histogram of the set of expenses from U.S. is nearly unimodal and symmetric; however, the distribution of the set of expenses from China is bimodal. According to Central Limit Theorem, we believe with bigger sample size, both distributions will be converging to normal distribution.

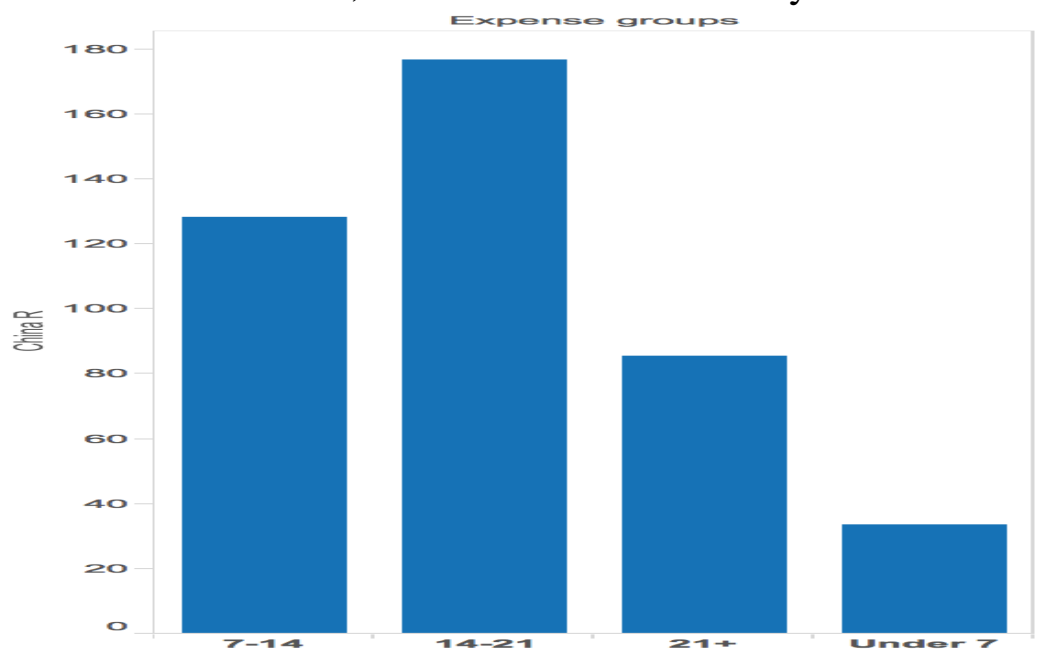


Distribution of expenses by American college students



Distribution of expenses by Chinese college students

Obviously I could choose to re-express the data by taking square root of all data I have. Then the distribution of expenses by Chinese college students is as follows, which satisfies the Nearly Normal Condition:



The assumptions are reasonable and the conditions are okay, so I'll use a Student's t-model to perform a **two-sample t-test**.

Significance Level

We set a significance level of 5%. That is, we reject null hypothesis if P-value is smaller than 5% (0.05).

Computation

From the data:

$$N_{US} = 22$$

$$\text{Mean}(\text{Exp}_{US}) = \$106.80$$

$$\text{Std}(\text{Exp}_{US}) = \$102.6666126$$

$$N_{China} = 35$$

$$\text{Mean}(\text{Exp}_{China}) = \$191.76$$

$$\text{Std}(\text{Exp}_{China}) = \$154.30748$$

For independent groups:

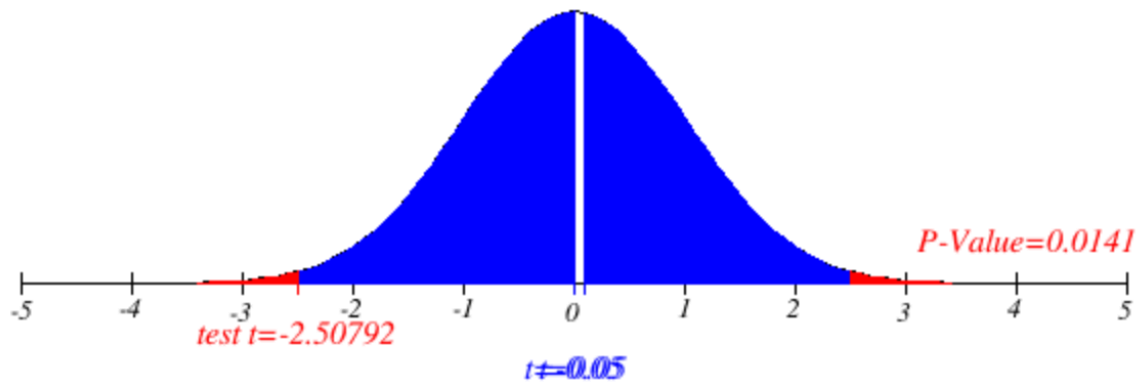
$$\text{SE}(\text{Mean}_{US} - \text{Mean}_{China}) = \text{Sqrt}(476.1106 + 671.5195264) = 33.87669$$

The observed difference is:

$$\text{Mean}(\text{Exp}_{US}) - \text{Mean}(\text{Exp}_{China}) = \$-84.96$$

$$t = (-84.96 - 0) / \text{SE}(\text{Mean}_{US} - \text{Mean}_{China}) = -84.96 / 33.87669 = -2.50792$$

The approximation formula gives 54.77 degrees of freedom. And I calculate P-value through computer. The Student-t distribution is plotted as below:



$P\text{-value} = 0.0141 < 0.05$

Conclusion and Interpretation

If there is no difference in the mean expenses, a difference this large would occur only 14 out of 1000 times. That's too rare to believe, so I reject the null hypothesis and conclude that Chinese college students are likely to spend more on cosmetics than American college students.

Limitations and Concerns

Conditions for Two-Sample T-Test

When I plotted out the distributions of two sets of expenses, clearly both distributions are not loosely, if not strictly, unimodal and symmetric, one with an outlier and another bimodal. An obvious solution is to sample the entire population through broader dissemination of designed surveys on not only social platforms but also offline in-person survey.

Convenience Sample

Truly I have to acknowledge that the sample I obtained is a convenience sample made up of my friends and friends' friends who are easy to reach. This sample is not representative of the entire population and thus generalization/ inference and conclusion is made conservatively. I find it obliging to describe the individuals who might be left out during the sampling process and individuals who might be overrepresented. The majority of my Chinese friends who I reached out are born and raised in Chinese middle class family whose parents pay over \$40,000 annually

for them to study abroad in the U.S. Many of the American college students I reached out are students studying at UT Austin with a few from Georgetown University where I had summer school. These are overrepresented people in the sample. The typical Chinese students that might be left out are those going to colleges in China with relatively low tuitions and those from working class. On the other hand, the American students that are not sufficiently sampled are those living in east and west coast and going to more prestigious private schools. If I obtained a more representative sample, I expect both distributions are nearly normal.

Analysis about the Conclusion

The conclusion that Chinese college students are likely to spend more on cosmetics than American college students is not surprising to me. From economic perspective, China is experiencing fast development and enjoying booming economy over recent ten years. Chinese people are no longer concerned about not having enough to eat and wear and began pursuing higher life quality through healthier lifestyle, more stylish food and restaurants, fashion and luxury, and definitely higher demand for beauty like cosmetic surgery and products. From cultural perspective, oriental culture is characterized with exquisite attention to details, and modern Chinese women (not only college students) buy various kinds of cosmetic products that correct the tiniest blemishes of face and body.

Obviously if a more representative population of American and Chinese college students is sampled, the distribution of the expenses on cosmetics by both groups will be nearly normal and inferences drawn will be more accurate.