

MATH 262 - Applied Statistics

Lab 6: Online Dating

Cesar O. Aguilar

In 2020, a research center conducted surveys pertaining to how relationships are affected by the internet. A portion of the surveys focused on online dating. The data is contained in the file `online_dating_data.csv`. The data includes the responses from 2,252 people age 18 and over. You will find five columns of data in the file. The data includes the following variables:

- **age** - the age of the interviewee, where 97 = 97 or older, 98 = don't know, and 99 = refused to answer.
- **gender** - the gender of the interviewee, where 1 = male, 2 = female.
- **onlineDating** - response to the question: "Have YOU, personally, ever used an online dating site?", where 1 = yes, 2 = no, 8 = don't know, 9 = refused to answer. If blank, assume the question did not apply to the interviewee.
- **metOnline** - response to the question: "Is your spouse/current partner someone you first met ONLINE or someone you first met OFFLINE?", where 1 = met online, 2 = met offline, 8 = don't know, 9 = refused to answer. If blank, assume the question did not apply to the interviewee.
- **desperate** - response to the question: "Do you agree or disagree with this statement: People who use online dating sites are desperate", where 1 = agree, 2 = disagree, 8 = don't know, 9 = refused to answer.
- **goodWay** - response to the question: "Do you agree or disagree with this statement: Online dating is a good way to meet people", where 1 = agree, 2 = disagree, 8 = don't know, 9 = refused to answer.

Answer the questions below using the following guidelines:

- (i) In all hypothesis test, use $\alpha = 0.05$ and report the p -value.

- (ii) **In each case, show what computations you performed to arrive at your answer. Simply providing a numerical answer without showing your work will result in a 0 for the question.**
- (iii) In each hypothesis test, make sure to filter out rows where the variable is either unknown or refused to answer. Consult the above description of the variables.
1. In a similar survey done in 2010, the mean age of people who reported having used an online dating site was 46 years old. Was the mean age of people who had used an online dating site different in 2020 than it was in 2010?
- (a) State the null and alternative hypotheses.
- (b) Perform the hypothesis test and state your conclusion regarding the null and alternative hypotheses. Use both a z -test and a t -test to see if there is a difference between the tests. Make sure to create and activate filters that include records where the **age** variable is known and **onlineDating == 1**.
- (c) What is your conclusion in the context of the problem?

2. In 2010, 2.6% of people who were married or in a serious relationship said they met online. Has this proportion increased in recent years?

(a) State the null and alternative hypotheses.

(b) Perform the hypothesis test and state your conclusion regarding the null and alternative hypotheses. To perform a test for proportion in `jamovi`, use the **2 Outcomes** test under the **Frequencies** menu. Make sure the appropriate filters are set active.

(c) What is your conclusion in the context of the problem?

3. In 2010, 29.5% of people agreed with the statement “People who use online dating are desperate.” Was this proportion different in 2020?

(a) State the null and alternative hypotheses.

(b) Perform the hypothesis test and state your conclusion regarding the null and alternative hypotheses. Make sure the appropriate filters are set active.

(c) What is your conclusion in the context of the problem?

4. Based on the 2020 survey, is the mean age of those who agreed with the statement “People who use online dating are desperate” different from those who disagreed?

(a) State the null and alternative hypotheses.

(b) Perform the hypothesis test and state your conclusion regarding the null and alternative hypotheses. Perform both a z -test and a t -test to see if there is any difference between the tests. Make sure the appropriate filters are set active.

(c) What is your conclusion in the context of the problem?

5. Based on the 2020 survey, do young people (age 18–29) tend to agree more with the statement “Online dating is a good way to meet people” than people who are over 29?

(a) State the null and alternative hypotheses.

(b) Perform the hypothesis test and state your conclusion regarding the null and alternative hypotheses. Make sure the appropriate filters are set active.

(c) What is your conclusion in the context of the problem?