

Name: Myra Rust

Date: 8 May 2020

Title: Predicting success rates of IVF

Section 1 – Week 9

- **Introduction**

Many women have difficulty becoming pregnant and look to assisted reproductive procedures to help with conceiving a child. In vitro fertilization (IVF) is one of the procedures available. During the IVF procedure, an egg is retrieved from the woman, fertilized in a laboratory, then the fertilized embryo is transferred to the woman's uterus. The IVF process is a very expensive process and therefore, someone looking to undergo the procedure would be interested in knowing success rates and factors that can affect success rates, in order to make decisions about the procedure and what is right for them.

- **Research questions**

1. What is the success rate of in vitro fertilization compared to other assisted reproductive procedures?
2. What is the success rate of in vitro fertilization by age?
3. What factors have the largest positive effect on the in vitro fertilization process?
4. What factors have negative effects on the in vitro fertilization process?
5. Does using multiple embryos increase your chances of success?
6. How much does using multiple embryos increase the likelihood of having twins or multiples?
7. Is it easier to treat male or female infertility?

- **Approach**

Question 1: Find the mean success rate for types of assisted reproduction procedures and compare them.

Question 2: Plot success rate of in vitro fertilization by age.

Question 3 -5: The success of the in vitro fertilization process is binary, either a woman becomes pregnant or does not. I will use logistic regression to analyze the variables affecting the success of the procedure.

Question 6: Research the correlation between the number of embryos and the number of children at birth. Create a logistic regression model that predicts the number of children at birth based on the number of embryos transferred.

- **How your approach addresses (fully or partially) the problem.**

Using descriptive statistics, I will be able to answer some basic questions of the success of in vitro fertilization in general. Using inferential statistics, I will draw conclusions about the success rate of the procedure based on the effects of particular variables. An individual should be able to review the information provided by the models and use it to help decide whether in vitro fertilization is a good choice for them.

- **Data**

<https://www.hfea.gov.uk/media/2667/ar-2015-2016-xlsb.xlsb>

This data has been collected in the UK specifically to improve fertility patient care by providing a data set for any researchers to conduct research on while maintaining patient confidentiality. The dataset is quite large at 158519 observations and 95 variables. I will need to reduce the size of this data set to make it usable because currently it takes quite a bit of time to load into RStudio.

- **Required Packages**

Readxl, broom, dplyr, car, ggplot2

- **Plots and Table Needs**

Question 1 could benefit from a plot of success rates by procedure type.

Question 2 will use a histogram of success rate by age group.

Question 6 could benefit from a visualization depicting the correlation between multiple embryos and multiple babies.

- **Questions for future steps**

- I need to become more familiar with my data set.
- I need to learn how to implement everything we've learned in this class without being told what technique to use, task to complete, or process to follow.