Name: Myra Rust   
Date: 24 May 2020   
Title: Predicting success rates of IVF

**Section 3 – Week 11**

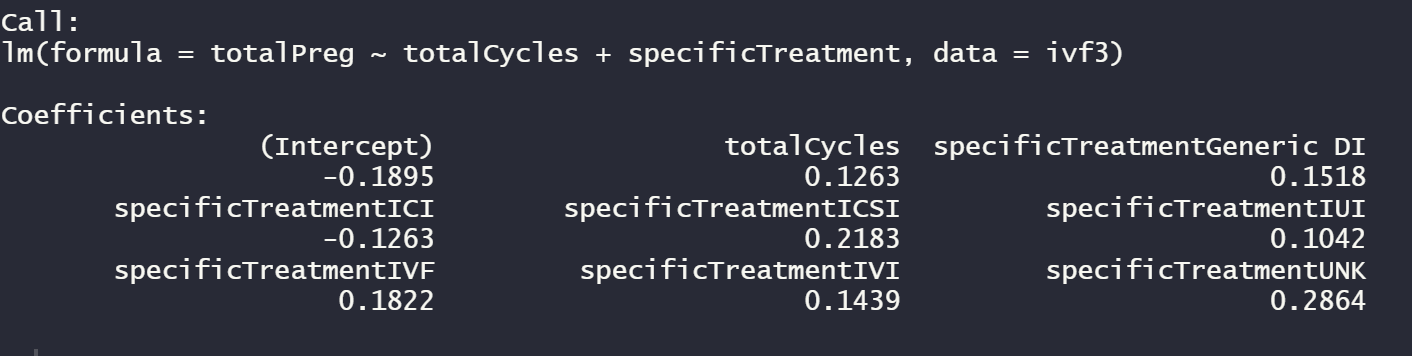
**What types of plots and tables will help you illustrate the findings to your questions?  (Ensure all graph/plots have axis titles, legends if necessary, scales where appropriate, geoms used, etc.).**

Research questions

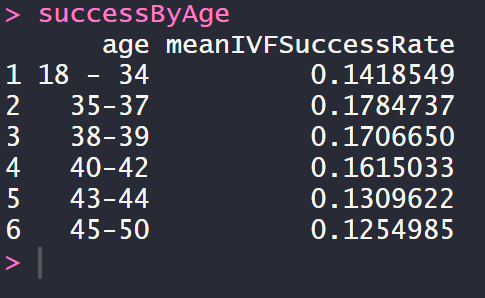
1. What is the success rate of IVF compared to other assisted reproductive procedures?

A screenshot of a video game

Description automatically generated



1. What is the success rate of IVF by age?

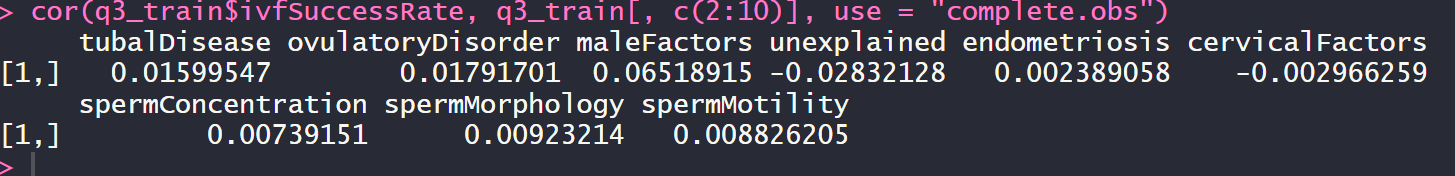


A picture containing photo, table, white

Description automatically generated

1. What factors have the largest positive/negative effect on the IVF process?

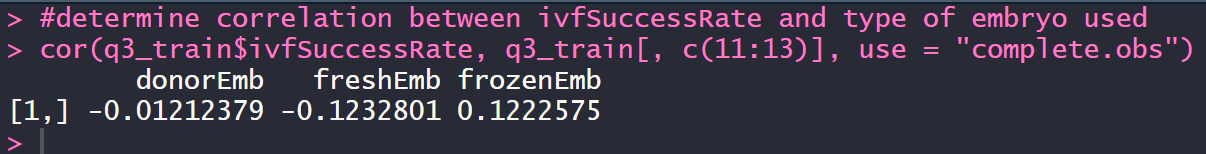
Correlation between ivfSuccessRate and infertility reasons:



A screenshot of a cell phone

Description automatically generated

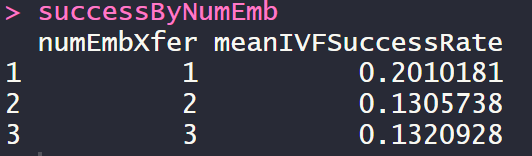
Correlation between ivfSuccessRate and type of embryo used:



A screenshot of a cell phone

Description automatically generated

1. Does using multiple embryos increase your chances of success?



A screenshot of a cell phone

Description automatically generated

1. How much does using multiple embryos increase the likelihood of having twins or multiples?

A screenshot of a cell phone

Description automatically generated

|  |  |
| --- | --- |
| Number of Embryos Transferred | Likelihood of having multiples |
| 1 | 1% |
| 2 | 22% |
| 3 | 13% |

**What topics have you learned in this class which most helped you answer your questions?**

With no previous experience in statistics or Language R, pretty much everything in this course was vital to being able to answer these questions. Since I learn best and retain the most by doing, I think that the data camp exercises were the best thing for me. However, while working with this data set, I realized a deficiency in my understanding of how to correctly prepare the data for this work. Originally, I thought my data set would be easy to work with, being that it was almost completely binary data. During this last week, I realized that was not true and I wish we had learned more about how to prepare and work with the data when it is not user friendly.

**What machine learning techniques do you plan on incorporating to answer your research questions?**

I plan to do regression testing to explain the data I have and predict future success rates. Most of my data is not linear, so this will be quite a task using different types of generalized linear models (glms). I currently don’t plan to conduct any unsupervised machine learning techniques; I’m planning to focus on regression and not bite of more than I can chew in a week.