**Memorandum**

To: NTRHD Intern

From: Dr. Brad Cannell, Director, NTRHD

Re: Analyzing a Diabetes Medical Device Study

**Instructions**

The NTRHD is doing an analysis of data acquired from a diabetes medical device study. The study took place across multiple clinics. Physicians collected data at each clinic, compiled the data, recorded the data electronically, and then emailed the data to us. We need you to import the raw data set into R and save it locally as an .Rds file that we will analyze later. Additionally, we want to compare this diabetes data to national diabetes data from the National Health and Nutrition Examination Survey (NHANES). So, I would like you to import, alter, and export that data as well.

**Task 1**. Click the link below to download the raw data file to your computer:

* [diabetes.xlsx](https://www.dropbox.com/s/0cizsa3ejiy3jy9/diabetes.xlsx?dl=1)
* Notice that this excel workbook contains two sheets. One sheet contains demographic information about the study participants. The other sheet contains study outcomes data.

**Task 2**. Import the diabetes study demographics data into R as a data frame named **diab\_demo.**

|  |
| --- |
| **Question:**  1. Which of the following code chunks will correctly import the diabetes study demographics data into R as a data frame named **diab\_demo**? |

**Task 3**. Coerce the **sex** variable to a factor. Name the factor version of **sex**, **sex\_f**. Possible values of **sex\_f** should include **female**, **male**, and **other** — in that order.

|  |
| --- |
| **Question:**  2. Fill in the following blanks to complete the code chunk that will coerce the **sex** variable to a factor with possible values of **female**, **male**, and **other** — in that order. |

**Task 4**. Use the str() function to view the structure of **diab\_demo.**

|  |
| --- |
| **Question:**  3. The **diab\_demo** data frame contains \_\_\_\_ rows and \_\_\_\_ columns. |

**Task 5**. Write code that instructs R to count the number of participants who fall into each category of **sex\_f**.

|  |
| --- |
| **Question:**  4. According to the analysis you completed in Task 5, \_\_\_\_ people reported that their **sex** was **other**. |

**Task 6**. Export **diab\_demo** to an .Rds file with the same name (i.e., **diab\_demo**).

|  |
| --- |
| **Question:**  5. Which of the following code chunks will correctly export the **diab\_demo** data frame as an .Rds file with the same name (i.e., **diab\_demo**)? |

**Task 7**. Import the diabetes study outcomes data into R as a data frame named **diab\_out**. Make sure to tell R how to correctly identify missing values in the raw data.

|  |
| --- |
| **Question:**  6. Which of the following code chunks will correctly import the diabetes study outcomes data into R as a data frame named **diab\_out**? |

**Task 8**. Coerce the **diabetes** variable to a factor variable. Name the factor version of **diabetes diabetes\_f**. Possible values that the **diabetes\_f** should include “No” and “Yes” — in that order.

|  |
| --- |
| **Question:**  7. Fill in the following blanks to complete the code chunk that will coerce the **diabetes** variable to a factor with zeros labeled as "**no**" and ones labeled as "**yes**" — in that order. |

**Task 9**. Use the str() function to view the structure of **diab\_out.**

**Task 10**. Export **diab\_out** to an .Rds file on your computer’s desktop.

|  |
| --- |
| **Question:**  8. Fill in the following blanks to complete the code chunk that will export **diab\_out** to an .Rds file with the same name on the computer’s desktop. |

**Task 11**. Import the 2015-2016 NHANES diabetes questionnaire data into R as a data frame named **diq\_i.**

**Task 12**. Use the glimpse() function to view the structure of **diq\_i**.

|  |
| --- |
| **Question:**  9. The **diq\_i** data frame contains \_\_\_\_ rows and \_\_\_\_ columns. |

**Task 13**. Coerce all of the variable names in **diq\_i** to lowercase.

|  |
| --- |
| **Question:**  10. Fill in the following blanks to complete the code chunk that will coerce all of the variable names in **diq\_i** to lowercase. |

**Task 14.** Print only the first 6 rows of **diq\_i** to your screen.

|  |
| --- |
| **Question:**  11. Fill in the following blanks to complete the code chunk that will print only the first 6 rows of **diq\_i** to the screen. |