## Questions for categories:

### General Syntax/Questions:

-Indexing (%in%, [1,], $ notation)

🡪 Subsetting, filters, !=

-Tidyverse data

-Vector, list, dataframe (combine to dataframe)

### Reporting:

-Reporting (TLF)

-Cum sum/Frequency Tables

-Chi Squares

### Data Manipulations:

* What is the output of this code segment?
  + Candidate 1 has the dosage of what?

### EDA:

-Histograms

-Normality assumptions

🡪 Count of size, distribution, missing values

-T-Tests

-Z

-ANOVA and significance

-ANCOVA and significance

-Sample Size and Power

### Modeling:

-General Regression Tests (GLM or LM)

-Mixed-Effect Modeling

-Survival

-Hierarchical Modeling

-Bayesian (Meta-Analysis)

- Posterior and prior

-Experimental Design

-Pharmacometric Modeling??

🡪 Adverse effects modeling

-NGS ??

### SAS to R:

* Xpt data output format
* Haven package use case

### CDISC Standards:

-AdAM datasets

-

Packages:

|  |  |
| --- | --- |
| **Package** | **Description** |
| dplyr | A fast, consistent tool for working with data frame like objects, both in memory and out of memory. |
| plyr | A set of tools that solves a common set of problems: you need to break a big problem down into manageable pieces, operate on each piece and then put all the pieces back together. For example, you might want to fit a model to each spatial location or time point in your study, summarise data by panels or collapse high-dimensional arrays to simpler summary statistics. |
| tidyr | Tools to help to create tidy data, where each column is a variable, each row is an observation, and each cell contains a single value. 'tidyr' contains tools for changing the shape (pivoting) and hierarchy (nesting and 'unnesting') of a dataset, turning deeply nested lists into rectangular data frames ('rectangling'), and extracting values out of string columns. It also includes tools for working with missing values (both implicit and explicit). |
| emmeans | Obtain estimated marginal means (EMMs) for many linear, generalized linear, and mixed models. Compute contrasts or linear functions of EMMs, trends, and comparisons of slopes. Plots and compact letter displays. |
| sqldf | The sqldf() function is typically passed a single argument which is an SQL select statement where the table names are ordinary R data frame names. sqldf() transparently sets up a database, imports the data frames into that database, performs the SQL select or other statement and returns the result using a heuristic to determine which class to assign to each column of the returned data frame. |
| flextable | Create pretty tables for 'HTML', 'Microsoft Word' and 'Microsoft PowerPoint' documents. Functions are provided to let users create tables, modify and format their content. It extends package 'officer' that does not contain any feature for customized tabular reporting and can be used within R markdown documents |
| officer | Access and manipulate 'Microsoft Word' and 'Microsoft PowerPoint' documents from R. The package focuses on tabular and graphical reporting from R; it also provides two functions that let users get document content into data objects. A set of functions lets add and remove images, tables and paragraphs of text in new or existing documents. When working with 'Word' documents, a cursor can be used to help insert or delete content at a specific location in the document. The package does not require any installation of Microsoft products to be able to write Microsoft files. |
| magrittr | Provides a mechanism for chaining commands with a new forward-pipe operator, %>%. This operator will forward a value, or the result of an expression, into the next function call/expression. |
|  |  |

<https://www.quanticate.com/blog/r-programming-in-clinical-trials>