MKTG 411 Marketing Analytics

Cheat Sheet

Data

R and Transform Data

`library()` load packages
`<-` assign objects
`read_csv()` import CSV files
`glimpse()` columns in a data frame
`filter()` observations using column values
`slice()` observations using their position in the data
`arrange()` observations in ascending or `desc()` order
`select()` variables
`mutate()` (i.e., recode or create) variables
`left_join()` two data frames and keep all columns on the "left"
`inner_join()` two data frames and keep all columns
`join_by()` helper function to specify the common ID for a join
`|>` together consecutive lines of code

Visualize Discrete Data

`count()` discrete data
`ggplot()` create a ggplot
`aes()` helper function for an aesthetic mapping
`geom_col()` column plot
`facet_wrap()` facet plot
`labs()` add labels
`scale_fill_manual()` specify the legend and color values for the `fill` variable
`unnest_tokens()` into a tidy text data frame
`drop_na()` drop rows with `NA`s
`anti_join()` remove rows that match
`fct_reorder()` to encode factor order

Visualize Continuous Data

`summarize()` data
`geom_histogram()` histogram
`geom_point()` scatterplot
`geom_smooth(method = "lm")` regression line
`geom_jitter()` jittered scatterplot
`group_by()` grouped summary
`slice_max()` slice top observations by group
`geom_line()` line plot
`mdy()` turn a month-day-year column into a `date`
`geom_boxplot()` boxplot
`geom_density()` desnity plot
`scale_color_manual()` specify the color values for the `color` variable
`theme()` and `theme_minimal()` to modify plot elements
`facet_grid()` facet plot with two variables

Quarto and Tidy Data

`pivot_longer()` turn column names into values
`c()` combine values into a vector (i.e., column)
`pivot_wider()` turn values into column names
`separate()` values in a column into two (or more) columns
`unite()` values in to or more columns into one column
`dmy()` turn a day-month-year column into a `date`
`scale_x_data()` modify the axis that has a `date` mapped to it
`class()` print out an object's data class
`tibble()` create a tibble (i.e., data frame)
`as_tibble()` coerce data into a tibble (i.e., data frame)
`as.*()` coerce vectors into specific data types
`geom_bar()` wrapper around `count()` and `geom_col()`
`fct_infreq()` to encode factor order based on frequency
`fct_rev()` reverse factor order

Database Queries

`here()` write a file path with reference to an RStudio Project
`install.packages()` on your hard drive *once*
`dbConnect()` create a database connection
`PostgreSQL()` protocol for a PostgreSQL database
`dbListTables()` print tables in a database
`tbl()` print a database table
`show_query()` show the SQL query
`collect()` and download a table from a database

`dbDisconnect()` disconnect from a database
`write_csv()` write a CSV file
`write_rds()` write a native R data file

Web Scraping and APIs

`read_html()` to scrape a webpage
`html_node()` extract data from a given HTML node
`html_text()` extract text from a given HTML node
`str_split()` split a string by a given regular expression
`str_view()` practice writing regular expressions
`unlist()` turn a list into a vector (i.e., column)
`qualtrics_api_credentials()` set credentials for Qualtrics API
`all_surveys()` print surveys in connection
`fetch_survey()` import a survey based on an ID

Dashboards

`selectInput()` to create input functions in a sidebar
`$` notation to select an element of a data frame or list
`renderPlot()` to create an interactive plot
`renderValueBox()` to create an interactive value box
`valueBox()` to place the value box in the reactive environment