Instructions and Data for 432 Quiz 1 for Spring 2025

Thomas E. Love, Ph.D.

2025-02-24

THESE ARE JUST THE INSTRUCTIONS

The complete PDF will be posted by noon on Friday 2025-02-28, at which point this document will be removed from the web site. Complete instructions are included in the complete PDF.

The purpose of this document is just to let people read over the instructions in advance, in case that's useful.

Links

All links relevant to this Quiz will be found on **Friday 2025-02-28** at https://thomaselove.github.io/432-2025/quiz1.html.

This will include details on how to access:

- the Main Document (this pdf) containing the instructions and questions
- the Google Form Answer Sheet, and
- the three data sets we are providing

If there are any changes to the Quiz after it is first posted, those changes will be posted to the Quiz 1 links page **and** you will be notified of any such changes in an email sent to you by Dr. Love.

Deadline

The deadline to complete your work and submit the Google Form Answer Sheet is **Wednesday, 2025-03-05 at Noon**. All of your answers must be submitted through the Google Form Answer Sheet found on the links page by the deadline, without exception. The form will close one hour after the deadline, and no extensions will be made available, so please do not wait until the last moment to submit. We will not accept any responses except through the Google Form.

Instructions

This PDF document is 26 pages long. There are 25 questions on this Quiz. It is to your advantage to answer all of the Questions. A blank response cannot possibly score better than an incorrect one, a guess might be correct (or at least partially correct), so you should definitely answer all of the questions.

The Google Form Answer Sheet

The Google Form Answer Sheet contains places to provide your responses to each question, and a final affirmation where you'll type in your name to tell us that you followed the rules for the Quiz. You must complete that affirmation and then submit your results. When you submit your results (in the same way you submit a Minute Paper) you will receive an email copy of your submission, with a link that will allow you to edit your results. The Answer Sheet works like a Minute Paper, in that you must be logged into Google via CWRU to access it.

If you wish to work on some of Quiz 1 and then return later, you can do this by [1] completing the final question (the affirmation) which asks you to type in your full name, and then [2] submitting the Quiz 1 Answer Sheet. You will then receive a link at your CWRU email which will allow you to return to the Quiz 1 Answer Sheet as often as you like without losing your progress.

The Data Sets

I have provided **three** data sets (called **dat_A.csv**, **dat_B.xlsx**, and **dat_C.Rds**) that are mentioned in the Quiz. They may be helpful to you, and are described later in this document.

What does the Quiz cover?

Quiz 1 includes material from the first 14 classes in 432, including all of Jeff Leek's *How to be a Modern Scientist*.

Scoring and Timing

Each question is worth 4 points, adding to a total of 100 points. If a question has two parts (a and b), then each part is worth 2 points. The questions are not in any particular order, and range in difficulty from "things Dr. Love expects everyone to get right" to "things that are deliberately tricky". Some questions will take more time than others to answer.

The Quiz is meant to take 4-6 hours to complete. I expect most students will take 3-8 hours, and some will take as little as 2 or as many as 10. Again, it is **not** a good idea to spend a long time on any one question. Dr. Love will grade the Quiz, and an answer sketch and grading rubric, along with students grades, will all be available by class time on Thursday 2025-03-06.

Getting Help

This is an open book, open notes quiz. You are welcome to consult the materials provided on the course website and that we've been reading in the class, but you are not allowed to discuss the questions on this quiz with anyone other than Professor Love and the teaching assistants. You will be required to complete a short affirmation that you have obeyed these rules as part of submitting the Quiz.

If you need clarification on a Quiz question, you have exactly two ways of getting help:

- You can ask your question via email to **Thomas dot Love at case dot edu**, or
- You can ask your question of Dr. Love directly in class 15 on 2025-03-04.

During the Quiz period (2024-02-28 through 2024-03-05) we will not answer questions about Quiz 1 except through the email and class session above.

- Please check your email **and** the Quiz 1 links page *before* submitting the final version of your Quiz, to see if Dr. Love has posted any changes there.
- Dr. Love promises to respond to all questions received before 9 AM on 2024-03-06 in a timely fashion.

When Should I ask for Help?

We recommend the following process.

- If you encounter a tough question, skip it, and build up your confidence by tackling other questions.
- When you return to the tough question, spend no more than 10-15 minutes on it. If you still don't have it, take a break (not just to do other questions) but an actual break.
- When you return to the question, it may be much clearer to you. If so, great. If not, spend 5-10 minutes on it, at most, and if you are still stuck, ask us for help.
- This is not to say that you cannot ask us sooner than this, but you should **never, ever** spend more than 20 minutes on any question without asking for help.

A few cautions about asking us questions

- Specific questions are more likely to get helpful answers.
- We will not review your code or your English for you.
- We will not tell you if your answer is correct, or if it is complete.

Writing Code into the Answer Sheet

Occasionally, we ask you to provide R code in your response. Do not include the `library()` command at any time for any of your code. Instead, assume in all questions that all relevant packages have been loaded in R.

Packages and Settings used by Dr. Love

This doesn't mean that I used all of these packages (I did not), or that you need to use all of these packages, nor does it mean that you are prevented from using other packages we've discussed in class to complete the Quiz, but all of the packages that I used in writing the Quiz and its answer sketch are listed below.

```
knitr::opts_chunk$set(comment = NA)

library(janitor)
library(naniar)
library(rms)
library(bestglm)
library(broom)
library(car)
library(caret)
library(cobalt)
library(cutpointr)
library(GGally)
library(glue)
library(gt)
library(haven)
library(mice)
library(mosaic)
library(olsrr)
library(patchwork)
library(readxl)
library(rsample)
library(tableone)
library(xfun)
library(yardstick)
library(easystats)
library(tidyverse)

theme_set(theme_lucid())
```

The `dat_A.csv` file

The subjects in the `dat_A.csv` file provided to you include responses from 557 subjects gathered by the National Health Interview Survey, specifically the child interviews completed in 2022[1]. The 8 variables included are specified in the table below.

Variable	Description
`subject_id`	child (subject) ID code
`sex`	child's sex: Male or Female
`adult_edu`	educational attainment of adults in child's family (5 levels)
`health_status`	child's health status (3 levels)
`race`	child's race/ethnicity (3 levels)
`violence`	1 = child has been a victim of or has witnessed violence, else 0
`age`	age of child in years (quantitative)
`sdq`	Strengths and Difficulties questionnaire total score (quantitative)

The `dat_B.xlsx` file

The `dat_B.xlsx` file provided to you includes information on 225 U.S. cities gathered by the Federal Bureau of Investigation's Quarterly Uniform Crime Report describing 2022 through voluntary city/agency submission to the National Incident-Based Reporting System (violent crime, property crime, population), the U.S. Census Bureau's "QuickFacts" page for each city (education, income, poverty), the U.S. Bureau of Labor (unemployment), and a set of regions selected from a map published by National Geographic[2]. The 8 variables included are specified in the table below.

Variable	Description
`city`	name of the city
`viol_rate`	violent crime rate per 100,000 people (2022)
`prop_high`	1 = city's property crime rate was above national average (1954.4 incidents per 100,000 people), 0 = city's property crime rate was below national average
`m_income`	median household income in 2022 dollars
`poverty`	poverty rate (as a percentage) of city residents, for 2022
`coll_grad`	% of city residents ages 25+ with a Bachelor's degree or higher education, 2018-2022

[1]Sarah Albalawi used a larger version of these data in her 432 Project A in 2024. Thanks, Sarah!
[2]Kathryn Menta gathered and combined these data, as part of her 432 Project A in 2024. Thanks, Kathryn!

Variable	Description
unemp	unemployment rate (as a percentage) of city residents, for December 2022
region	region of the country (3 levels: Eastern, Middle and Western)

The dat_C.Rds file

The dat_C.Rds file provided to you includes information on 1200 female subjects from the Behavioral Risk Factor Surveillance System (BRFSS) data for 2022, who are ages 25-65, live in Virginia, and have not had a hysterectomy[3]. The 9 variables included are specified in the table below.

Variable	Description
ID	subject identifier
bmi	subject's body mass index, in kg/m^2
cv_scr	1 if subject is up to date on cervical cancer screening per USPSTF and ACOG guidelines, 0 if subject is not up to date
race_eth	subject's self-defined race-ethnicity (5 categories)
physhlth	# of days in the past 30 when subject's physical health was not good
arthritis	does subject have arthritis? (Yes or No)
exercise	did subject participate in physical activity for exercise out of work in the past 30 days? (Yes or No)
c_assault	did subject experience sexual assault (forced touching or intercourse) as an adverse childhood experience (1 = Yes, 0 = No)
sex_ori	subject's sexual orientation (2 levels: "Straight", "Not_Straight")

Session Information

```
session_info()
```

```
R version 4.4.2 (2024-10-31 ucrt)
Platform: x86_64-w64-mingw32/x64
Running under: Windows 11 x64 (build 26100)

Locale:
```

[3]Liz Stanley gathered and combined these data, as part of her 432 Project A in 2024. Thanks, Liz!

```
  LC_COLLATE=English_United States.utf8
  LC_CTYPE=English_United States.utf8
  LC_MONETARY=English_United States.utf8
  LC_NUMERIC=C
  LC_TIME=English_United States.utf8

Package version:
  abind_1.4-8           askpass_1.2.1         backports_1.5.0
  base64enc_0.1-3       bayestestR_0.15.2     bestglm_0.37.3
  bigD_0.3.0            bit_4.5.0.1           bit64_4.6.0.1
  bitops_1.0.9          blob_1.2.4            boot_1.3-31
  broom_1.0.7           bslib_0.9.0           cachem_1.1.0
  callr_3.7.6           car_3.1-3             carData_3.0-5
  caret_7.0-1           cellranger_1.1.0      checkmate_2.3.2
  chk_0.10.0            class_7.3-22          cli_3.6.4
  clipr_0.8.0           clock_0.7.2           cluster_2.1.6
  cobalt_4.5.5          coda_0.19-4.1         codetools_0.2-20
  colorspace_2.1-1      commonmark_1.9.2      compiler_4.4.2
  conflicted_1.2.0      correlation_0.8.6     cowplot_1.1.3
  cpp11_0.5.1           crayon_1.5.3          curl_6.2.1
  cutpointr_1.2.0       data.table_1.16.4     datasets_4.4.2
  datawizard_1.0.0      DBI_1.2.3             dbplyr_2.5.0
  Deriv_4.1.6           diagram_1.6.5         digest_0.6.37
  doBy_4.6.25           dplyr_1.1.4           dtplyr_1.3.1
  e1071_1.7.16          easystats_0.7.4       effectsize_1.0.0
  emmeans_1.10.7        estimability_1.5.1    evaluate_1.0.3
  fansi_1.0.6           farver_2.1.2          fastmap_1.2.0
  fontawesome_0.5.3     forcats_1.0.0         foreach_1.5.2
  foreign_0.8-88        Formula_1.2-5         fs_1.6.5
  furrr_0.3.1           future_1.34.0         future.apply_1.11.3
  gargle_1.5.2          gdata_3.0.1           generics_0.1.3
  GGally_2.2.1          ggformula_0.12.0      ggplot2_3.5.1
  ggridges_0.5.6        ggstats_0.8.0         glmnet_4.1-8
  globals_0.16.3        glue_1.8.0            gmodels_2.19.1
  goftest_1.2-3         googledrive_2.1.1     googlesheets4_1.1.1
  gower_1.0.2           graphics_4.4.2        grDevices_4.4.2
  grid_4.4.2            gridExtra_2.3         grpreg_3.5.0
  gt_0.11.1             gtable_0.3.6          gtools_3.9.5
  hardhat_1.4.1         haven_2.5.4           highr_0.11
  Hmisc_5.2-2           hms_1.1.3             htmlTable_2.4.3
  htmltools_0.5.8.1     htmlwidgets_1.6.4     httpuv_1.6.15
  httr_1.4.7            ids_1.0.1             insight_1.0.2
  ipred_0.9-15          isoband_0.2.7         iterators_1.0.14
```

janitor_2.2.1 jomo_2.7-6 jquerylib_0.1.4
jsonlite_1.9.0 juicyjuice_0.1.0 KernSmooth_2.23.24
knitr_1.49 labeling_0.4.3 labelled_2.14.0
later_1.4.1 lattice_0.22-6 lava_1.8.1
leaps_3.2 lifecycle_1.0.4 listenv_0.9.1
lme4_1.1-36 lubridate_1.9.4 magrittr_2.0.3
markdown_1.13 MASS_7.3-64 Matrix_1.7-1
MatrixModels_0.5-3 memoise_2.0.1 methods_4.4.2
mgcv_1.9.1 mice_3.17.0 microbenchmark_1.5.0
mime_0.12 minqa_1.2.8 mitml_0.4-5
mitools_2.4 modelbased_0.9.0 ModelMetrics_1.2.2.2
modelr_0.1.11 mosaic_1.9.1 mosaicCore_0.9.4.0
mosaicData_0.20.4 multcomp_1.4-28 munsell_0.5.1
mvtnorm_1.3-3 naniar_1.1.0 nlme_3.1-166
nloptr_2.1.1 nnet_7.3-20 norm_1.0.11.1
nortest_1.0-4 numDeriv_2016.8.1.1 olsrr_0.6.1
openssl_2.3.2 ordinal_2023.12.4.1 pan_1.9
parallel_4.4.2 parallelly_1.42.0 parameters_0.24.1
patchwork_1.3.0 pbkrtest_0.5.3 performance_0.13.0
pillar_1.10.1 pkgconfig_2.0.3 pls_2.8-5
plyr_1.8.9 polspline_1.1.25 prettyunits_1.2.0
pROC_1.18.5 processx_3.8.6 prodlim_2024.06.25
progress_1.2.3 progressr_0.15.1 promises_1.3.2
proxy_0.4.27 ps_1.9.0 purrr_1.0.4
quantreg_6.00 R6_2.6.1 ragg_1.3.3
rappdirs_0.3.3 rbibutils_2.3 RColorBrewer_1.1-3
Rcpp_1.0.14 RcppArmadillo_14.2.3.1 RcppEigen_0.3.4.0.2
Rdpack_2.6.2 reactable_0.4.4 reactR_0.6.1
readr_2.1.5 readxl_1.4.3 recipes_1.1.1
reformulas_0.4.0 rematch_2.0.0 rematch2_2.1.2
report_0.6.1 reprex_2.1.1 reshape2_1.4.4
rlang_1.1.5 rmarkdown_2.29 rms_7.0-0
rpart_4.1.24 rsample_1.2.1 rstudioapi_0.17.1
rvest_1.0.4 sandwich_3.1-1 sass_0.4.9
scales_1.3.0 see_0.10.0 selectr_0.4.2
shape_1.4.6.1 shiny_1.10.0 slider_0.3.2
snakecase_0.11.1 sourcetools_0.1.7.1 SparseM_1.84-2
sparsevctrs_0.2.0 splines_4.4.2 SQUAREM_2021.1
stats_4.4.2 stats4_4.4.2 stringi_1.8.4
stringr_1.5.1 survey_4.4-2 survival_3.8-3
sys_3.4.3 systemfonts_1.2.1 tableone_0.13.2
textshaping_1.0.0 TH.data_1.1-3 tibble_3.2.1
tidyr_1.3.1 tidyselect_1.2.1 tidyverse_2.0.0

timechange_0.3.0 timeDate_4041.110 tinytex_0.55
tools_4.4.2 tzdb_0.4.0 ucminf_1.2.2
UpSetR_1.4.0 utf8_1.2.4 utils_4.4.2
uuid_1.2.1 V8_6.0.1 vctrs_0.6.5
viridis_0.6.5 viridisLite_0.4.2 visdat_0.6.0
vroom_1.6.5 warp_0.2.1 withr_3.0.2
xfun_0.51 xml2_1.3.6 xplorerr_0.2.0
xtable_1.8-4 yaml_2.3.10 yardstick_1.3.2
zoo_1.8-13