**RESEARCH AND SURVEY STATISTICS – STA3022F**

**COMPUTER PRACTICAL 3**

**FACTOR ANALYSIS**

1. Save the following from VULA:

prac3.R

airman.csv

donor.csv

derived.csv

1. Open the script file **prac3.R**.

From the top menu, select File, Open File

Select the file prac3.R which you ***saved*** from Vula

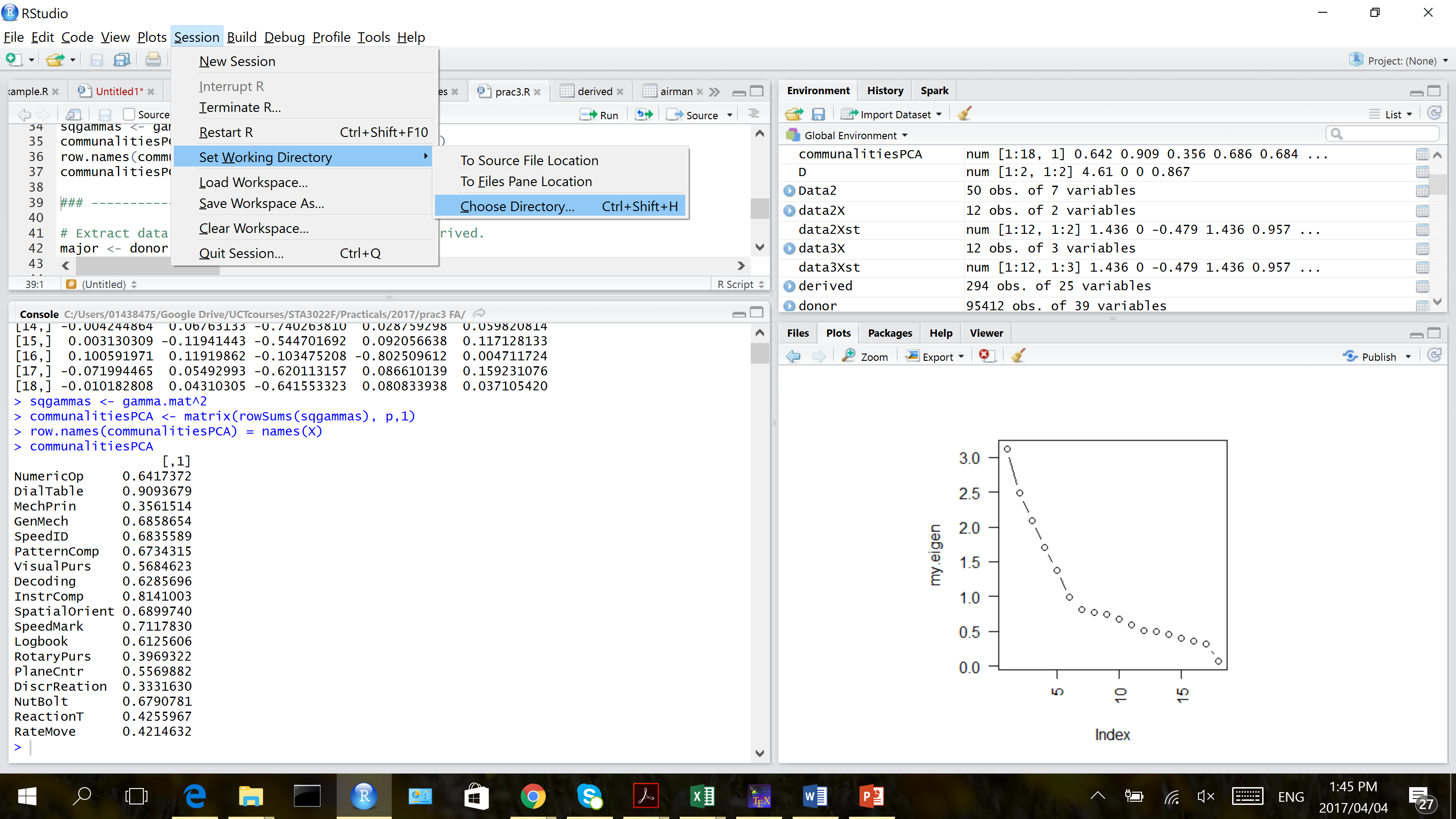
Click Open

The script file appears in the Window in the top left.

We will execute the commands step by step by selecting (highlighting) the relevant lines in the script file and clicking on Run at the top right of this window.

1. Load the airman.csv data into R.

Navigate to Session, Set Working Directory, Choose Directory. Here Navigate to where you saved your csv files.



Re-active the script file by clicking on prac3.R.

1. Example FA with airman data set

The data set is similar to that of Fleishman and Hempel[[1]](#footnote-1) (1954), used as an example in Maxwell[[2]](#footnote-2) (1977). The data simulated here are for 18 tests administered to 791 airmen.

Numerical Operations

Dial and Table Reading

Mechanical Principles

General Mechanics

Speed of Identification

Pattern Comprehension

Visual Pursuit

Decoding

Instrument Comprehension

Spatial Orientation

Speed of Marking

Log Book Accuracy

Rotary Pursuit

Plane Control

Discrimination Reaction Time

Nut and Bolt

Reaction Time

Rate of Movement

Import the data set airman.csv into R. Construct data matrix of the 8 tests and 791 participants.

Run lines 6 to 11.

1. Create a scree plot by selecting lines 15 to 17 and click run.
2. We will extract five factors here. Run line 20.
3. Perform FA with maximum likelihood estimation and varimax rotation and inspect the loadings.

Select line 23-28 and Run.

The following factors are identified:

ML2: Speed and accuracy of clerical operations

NumOp, SpeedID, Decoding, SpeedMark, LogbookAcc

ML1: Comprehension of spatial relations

DialTable, Pattern, Visual, Spatial

ML3: Facility with numbers

NumOp, DialTable, Decoding, Instr

ML5: Rate of movement

Rotary, PlaneCtrl, DiscReation, ReactionT, MoveRate

ML4: Mechanical aptitude and expertise

Mech, GenMech, NutBolt

Perform FA with PCA estimation and inspect the factor loadings. Estimate the factor loadings with PCA, select lines 31-38 and run.

The following factors are identified:

PA1: Speed and accuracy of clerical operations

NumOp, SpeedID, Decoding, SpeedMark, LogbookAcc

PA2: Comprehension of spatial relations

DialTable, Pattern, Visual, Spatial

PA5: Facility with numbers

NumOp, DialTable, Decoding, Instr

PA3: Rate of movement

Rotary, PlaneCtrl, DiscrReation, ReactionT, MoveRate

PA4: Mechanical aptitude and expertise

Mech, GenMech, NutBolt

Estimate the factor loadings with PCA, select lines 31-38 and run

Communalities for each observed variable:

|  |  |  |  |
| --- | --- | --- | --- |
|  | MLE | PCA | Unique (Specific Var) MLE |
| NumericOp | 0.49 | 0.64 | 0.51 |
| DialTable | 1.00 | 0.91 | 0.01 |
| MechPrin | 0.13 | 0.36 | 0.87 |
| GenMech | 0.56 | 0.69 | 0.44 |
| SpeedID | 0.64 | 0.68 | 0.36 |
| PatternComp | 0.63 | 0.67 | 0.37 |
| VisualPurs | 0.34 | 0.57 | 0.66 |
| Decoding | 0.64 | 0.63 | 0.36 |
| InstrComp | 0.56 | 0.81 | 0.44 |
| SpatialOrient | 0.56 | 0.69 | 0.44 |
| SpeedMark | 0.57 | 0.71 | 0.43 |
| Logbook | 0.43 | 0.61 | 0.57 |
| RotaryPurs | 0.26 | 0.40 | 0.74 |
| PlaneCntr | 0.48 | 0.56 | 0.52 |
| DiscrReation | 0.19 | 0.33 | 0.81 |
| NutBolt | 0.50 | 0.68 | 0.50 |
| ReactionT | 0.23 | 0.43 | 0.77 |
| RateMove | 0.24 | 0.42 | 0.76 |

1. Donor data analysis for quiz:

EXTRACT DONOR and DERIVED DATASETS: SELECT LINES 42-43:

Detailed information is available for major donors only. The frequency and amount donated is found in the derived data frame. Perform a Factor Analysis on the data for major donors to determine the main features of donor behaviour.

HIT Indicates total number of known times the donor has responded to a mail order offer other than PVA.

freqxx Frequency of donations during the year 19xx

freqTy Frequency of donations for each donation type

T1: mailings had labels only

T2: mailings are calendars with stickers but do not have labels

T3: mailings are blank cards that fold into thirds with labels

T4: mailings are blank cards with labels

T5: mailings have thank you printed on the outside with labels

T6: mailings are general greeting cards (an assortment of birthday, sympathy, blank, & get well) with labels

T7: mailings are Christmas cards with labels

T8: mailings have labels and a note pad

amountxx Dollar amount donated during the year 19xx

amountTy Dollar amount donated for each donation type

CARDPROM Lifetime number of card promotions received to date.

NUMPROM Lifetime number of promotions received to date.

CARDPM12 Number of card promotions received in the last 12 months.

NUMPRM12 Number of promotions received in the last 12 months.

RAMNTALL Dollar amount of lifetime donations to date

NGIFTALL Number of donations to date

CARDGIFT Number of donations to card promotions to date

MINRAMNT Dollar amount of smallest donation to date

MAXRAMNT Dollar amount of largest donation to date

LASTGIFT Dollar amount of most recent donation

AVGGIFT Average dollar amount of donations to date

The code and comments in the script file, lines 45-53 will assist you.

First run 45-53 and go to line 8 and run the rest of the codes similarly.

Remember to carefully consider whether you need to standardise your data. Complete the Quiz: Prac 3 FA.

1. Fleisman, E.A and Hempel, W.E. 1954. Changes in factor structure of a complex psychomotor test as a function of Practice. *Psychometrika*, **19**, 239-252. [↑](#footnote-ref-1)
2. Maxwell, A.E. 1977. Multivariate Analysis in Behavioural Research. Chapman and Hall: London. [↑](#footnote-ref-2)