## Multivariate Normality

### Correlation

Mvn1 & mvn2 = 0.27

Mvn1 & mvn3 = 0.54

Everything else is close to 0

### Outliers

#### Multivariate Outliers

A screenshot of a computer

Description automatically generated

Probably need to be removed, check again after transformations

#### Univariate Outliers

No information that these outliers were the result of coding/entry errors, instrument malfunction, or that they belong to another population than the one being studied. As such, we have to assume that these observations reflect legitimate variability in the population and should be included in the analysis. So Transformations should be considered (Raykov & Marcoulides, 2008)

* Mvn1:
  + A screenshot of a computer

    Description automatically generated
  + A screenshot of a computer

    Description automatically generated
* Mvn2:
  + Positively skewed because of crazy positive outliers
  + A black screen with white text

    Description automatically generated
  + A screenshot of a computer program

    Description automatically generated
* Mvn3
  + Negatively skewed with a few negative outliers (none extreme)
  + A screenshot of a computer

    Description automatically generated
  + A screenshot of a computer program

    Description automatically generated
* Mvn4:
  + Bimodal distribution
  + No outliers (0% > 1.5 IQR above Q3 or > 3.29 z)
* Mvn5:
  + Bimodal distribution around 0 and 100, suggesting floor and ceiling effects
  + No outliers (0% > 1.5 IQR above Q3 or > 3.29 z)

### Univariate Normality

* None of mvn1-5 respect univariate normality, both according to the Shapiro-Wilk test (explain assumptions) and from visual inspection of the QQ plots.
  + Checked with MVN package (Korkmaz et al., 2014)
* Distributional characteristics:
  + Mvn1 positively skewed
  + Mvn2 is positively skewed and has a very large outlier -> check distribution after outlier removed
    - Skew and kurtosis are very large
  + Mvn3 is negatively skewed
  + Mvn4 is bimodal
  + Mvn5 is bimodal -> probably as a result of ceiling and floor effect.
* Skewness and kurtosis were calculated using Fisher’s formulas (Cain et al., 2017) implemented in the psych package (Revelle, 2024), although the results of all common formulas are indistinguishable in large samples such as this one (Joanes & Gill, 1998).

### Transformations

#### MVN1

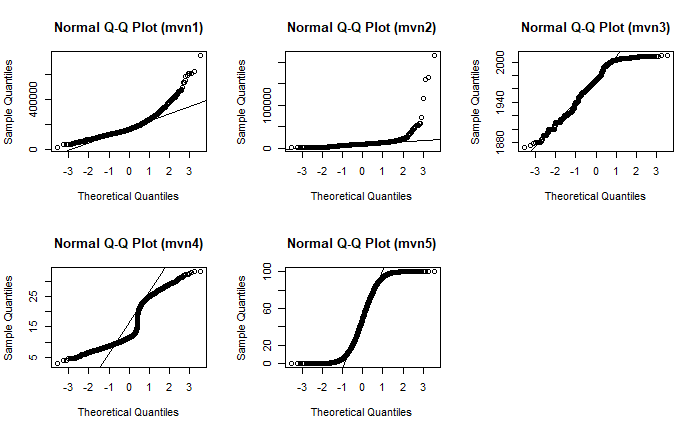
* Positive skew and fat tails
  + Skew = 1.76, kurtosis = 5.43
* Minimum = 12,789
  + Needs to be adjusted to minimum = 1 by subtracting 12,788
* Try square root transformation first, still left too much positive skew even after removing outliers past z = 3.29
  + Skew = 0.61, kurtosis = 0.61
* ln transformation
  + skew = 0.11
  + kurtosis 0.53
  + sw = 0.99
    - “To perform the S–W test for normality, assume that the sample is composed of *n* independent and identically distributed observations (*x*1,*x*2,…*xn*) from a [normal distribution](https://www.sciencedirect.com/topics/mathematics/gaussian-distribution) with unspecified mean and variance.” (Stephen W. Looney, Joseph L. Hagan, in [Essential Statistical Methods for Medical Statistics](https://www.sciencedirect.com/book/9780444537379/essential-statistical-methods-for-medical-statistics), 2011)
  + after removing outliers past z = 3.29

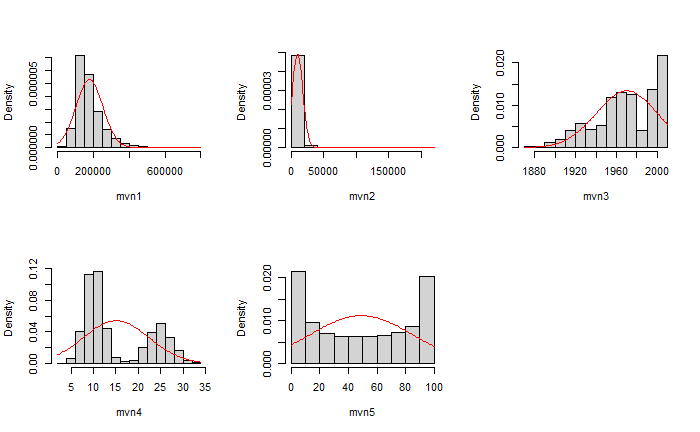
#### MVN2

#### MVN3

#### MVN4

#### MVN5





## Group Comparison

### Missing Data

## Preparing for Regression

## Repeated-Measures Screening

## References

William Revelle (2024). \_psych: Procedures for Psychological, Psychometric, and Personality Research\_. Northwestern University, Evanston, Illinois. R package version 2.4.6, <https://CRAN.R-project.org/package=psych>.

Korkmaz S, Goksuluk D, Zararsiz G. MVN: An R Package for Assessing Multivariate Normality. The R Journal. 2014 6(2):151-162.

**Cain, M. K., Zhang, Z., & Yuan, K.-H. (2017). Univariate and multivariate skewness and kurtosis for measuring nonnormality: Prevalence, influence and estimation. Behavior Research Methods, 49(5), 1716–1735.** [**https://doi.org/10.3758/s13428-016-0814-1**](https://doi.org/10.3758/s13428-016-0814-1)

**Joanes, D. N., & Gill, C. A. (1998). Comparing Measures of Sample Skewness and Kurtosis. *Journal of the Royal Statistical Society. Series D (The Statistician)*, *47*(1), 183–189.**

**Raykov, T., & Marcoulides, G. A. (2008). *An Introduction to Applied Multivariate Analysis* (1st edition). Routledge.**