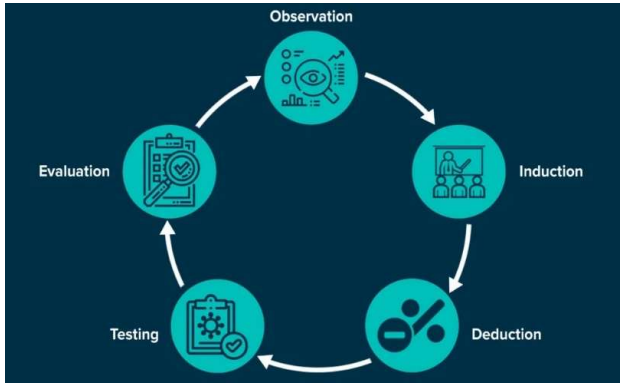
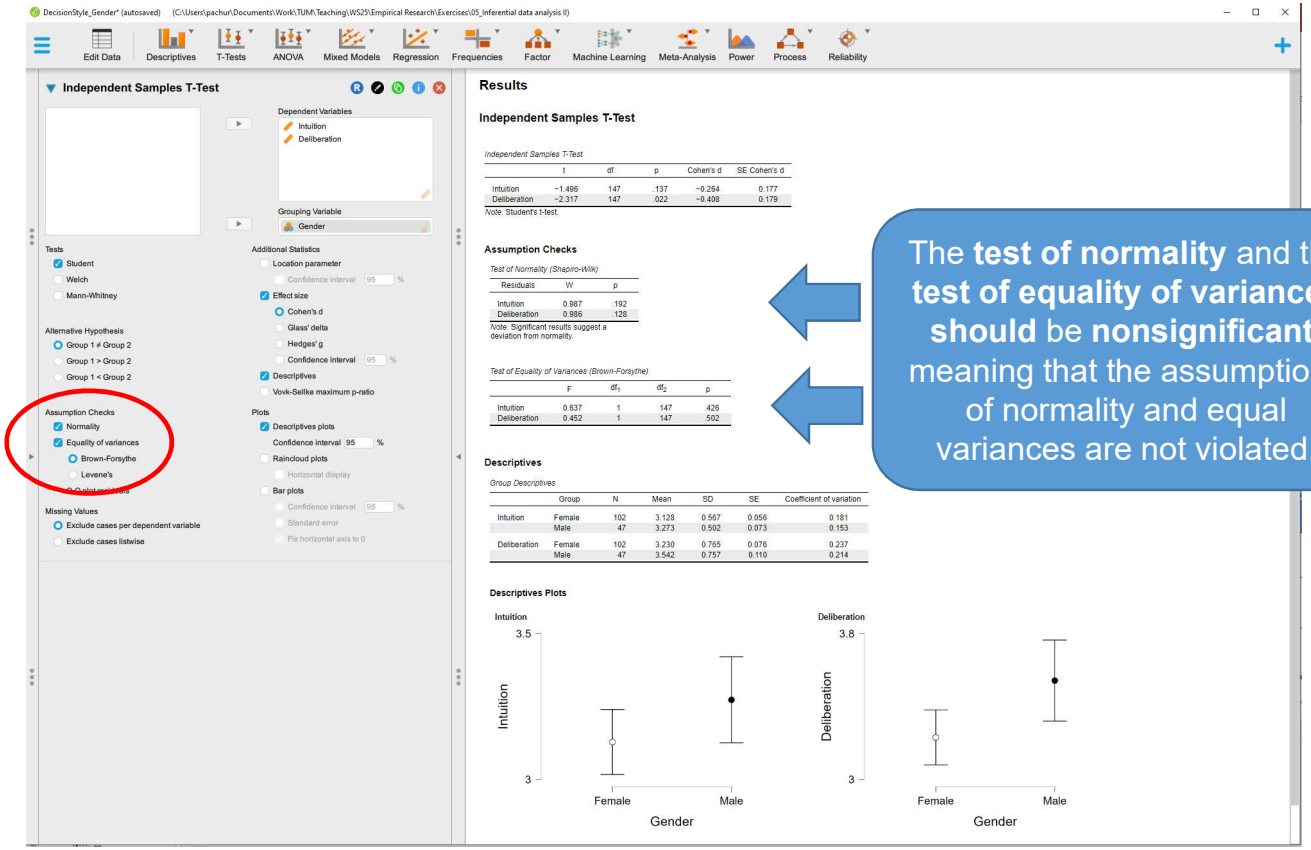
Empirical research in management and economics
Exercise

Thorsten Pachur, Linus Hof, Rebecca West, Sebastian Hellmann, Nuno Busch

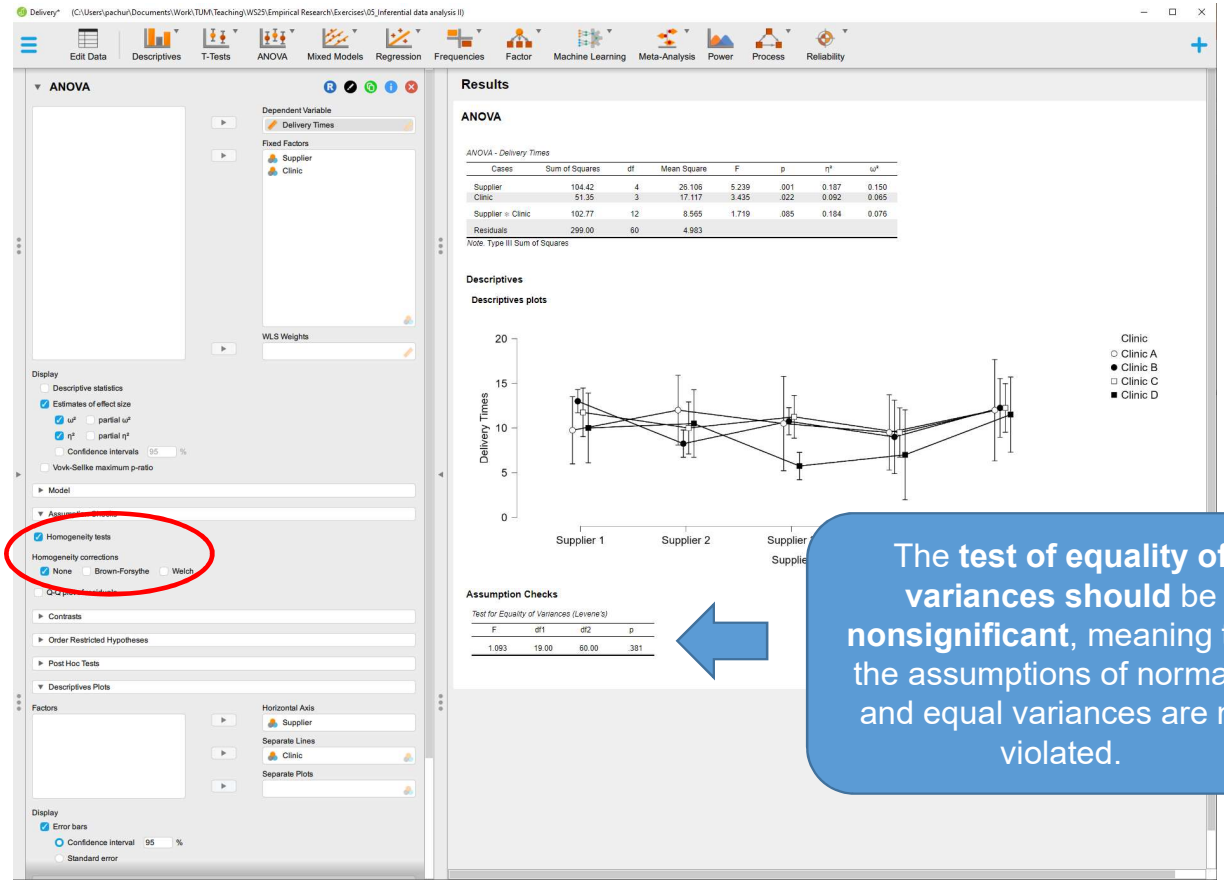
Technical University of Munich
School of Management
Chair of Behavioral Research Methods



Checking normality assumptions: t-test



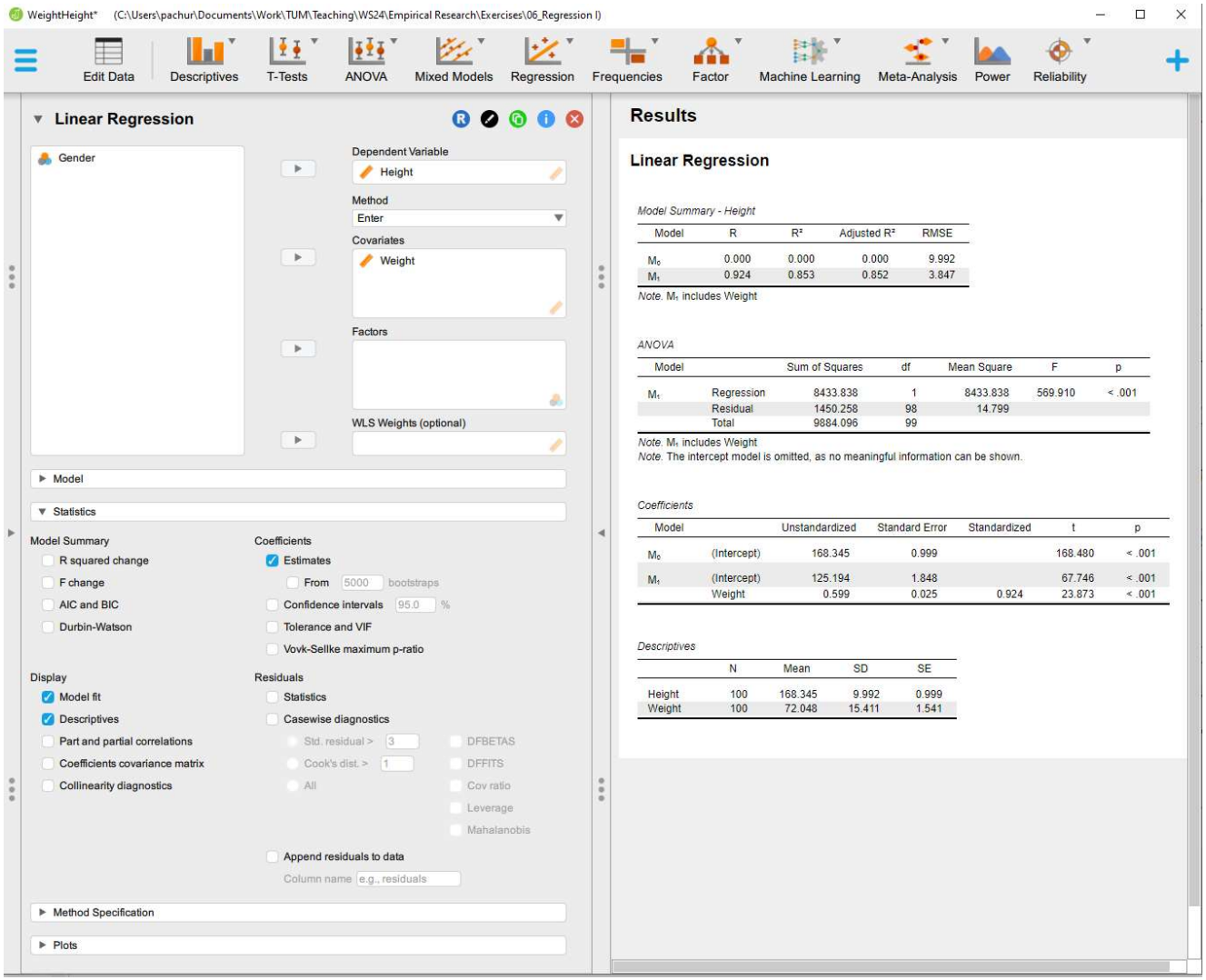
Checking normality assumptions: ANOVA

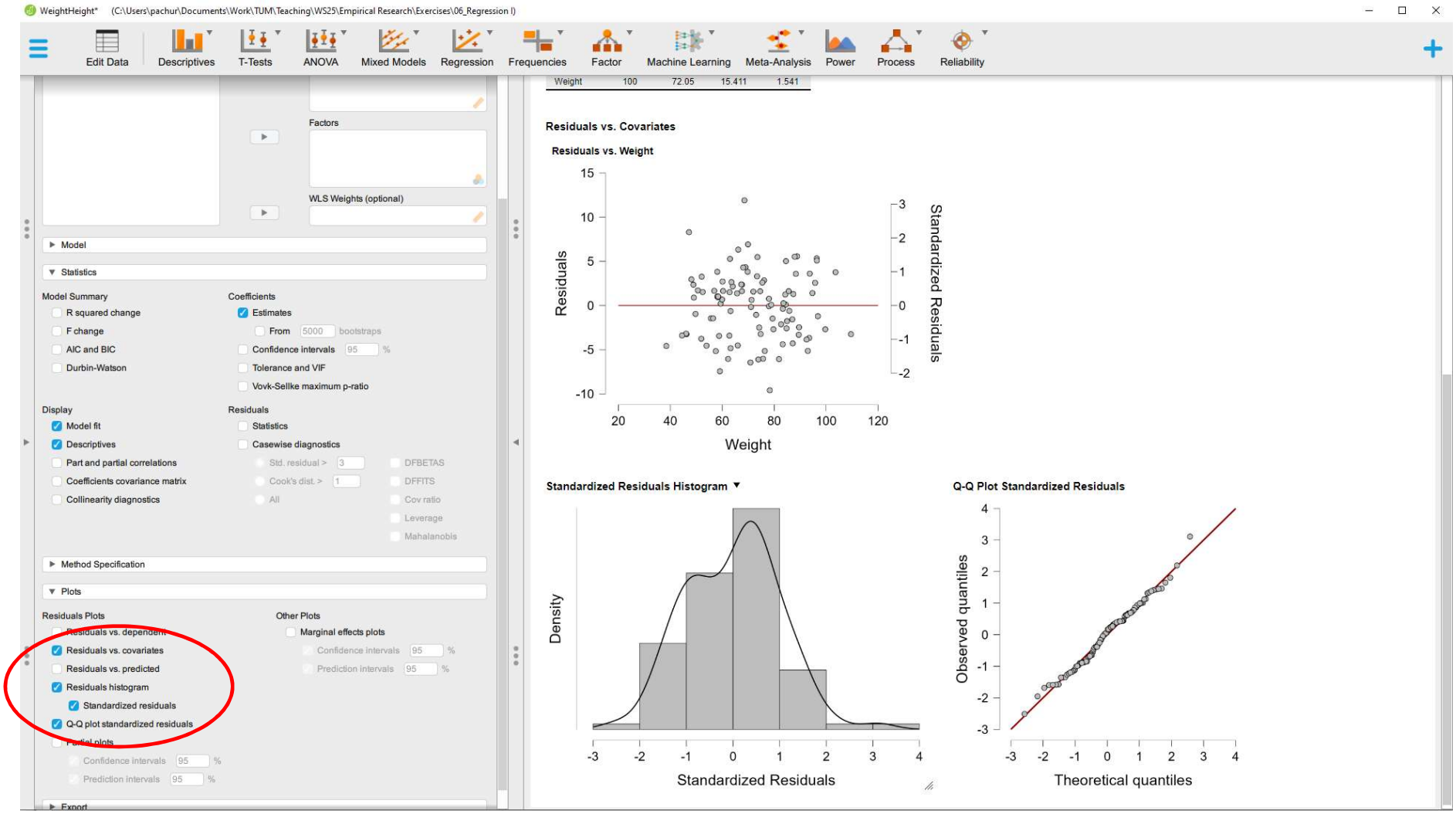


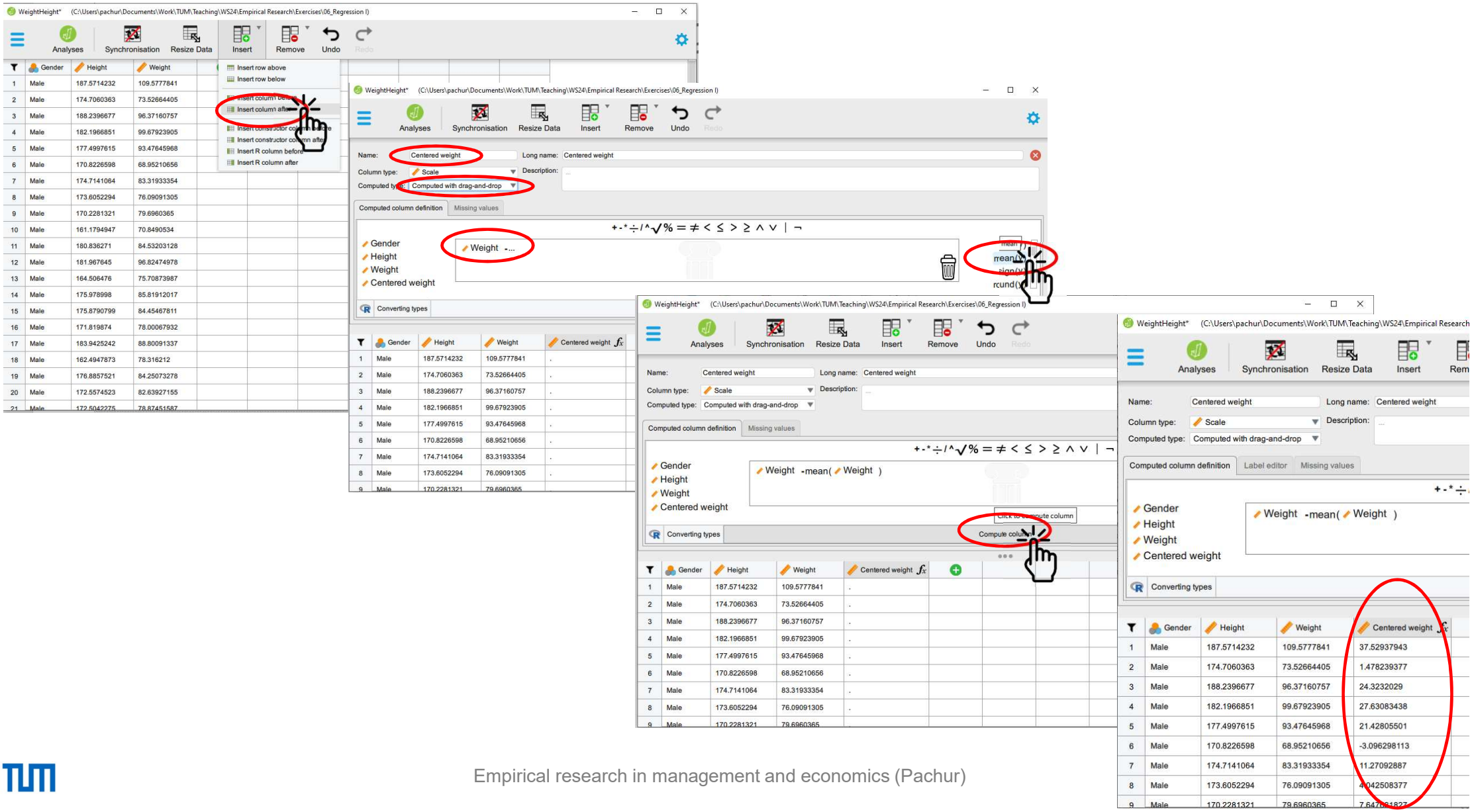
The **test of equality of variances should** be **nonsignificant**, meaning that the assumptions of normality and equal variances are not violated.

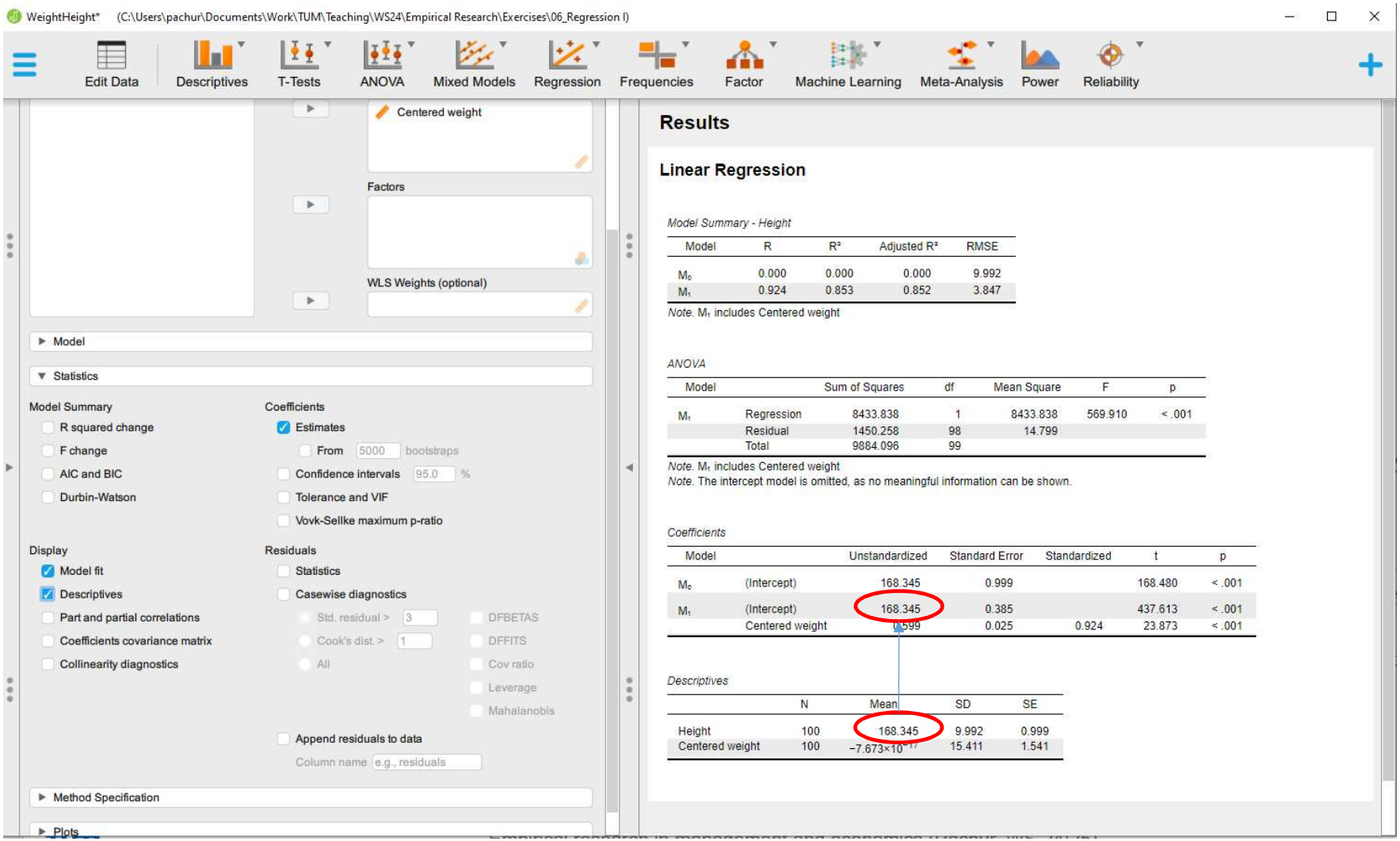
Exercise I

- Simple regression in JASP
- Open data file “WeightHeight.csv“
- Conduct a simple regression analysis to predict height from weight
  - Regression coefficient, significance test, R^2
  - Test assumptions: linearity, homoscedasticity, and normality
- Run the analysis when the predictor is centered (note: $X_{centered} = X - \bar{X}$)

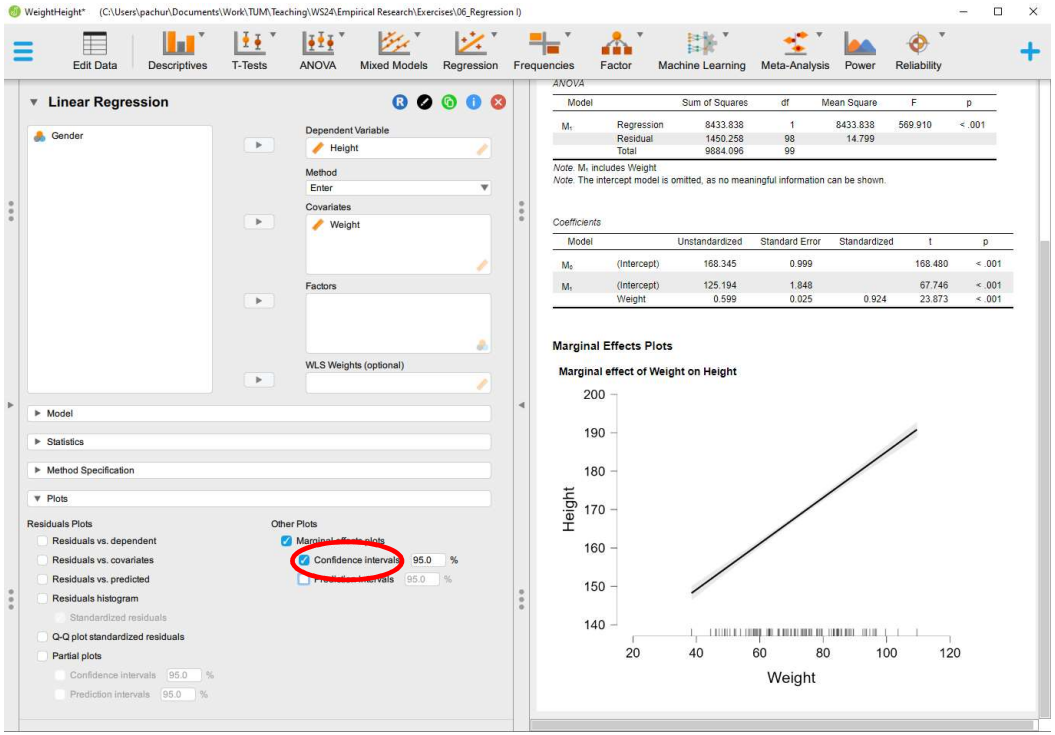

WeightHeight* (C:\Users\pachur\Documents\Work\TUM\Teaching\WS24\Empirical Research\Exercises\06_Regression I)
Edit Data
Descriptives
T-Tests
ANOVA
Mixed Models
Regression
Frequencies
Factor
Machine Learning
Meta-Analysis
Power
Reliability
Linear Regression
Gender
Dependent Variable
Height
Method
Enter
Covariates
Weight
Factors
WLS Weights (optional)
Model
Statistics
Model Summary
R squared change
F change
AIC and BIC
Durbin-Watson
Coefficients
Estimates
From 5000 bootstraps
Confidence intervals 95.0 %
Tolerance and VIF
Vovk-Sellke maximum p-ratio
Display
Model fit
Descriptives
Part and partial correlations
Coefficients covariance matrix
Collinearity diagnostics
Residuals
Statistics
Casewise diagnostics
Std. residual > 3
Cook's dist. > 1
All
DFBETAS
DFFITS
Cov ratio
Leverage
Mahalanobis
Append residuals to data
Column name e.g., residuals
Method Specification
Plots
Results
Linear Regression
Model Summary - Height
Model R R² Adjusted R² RMSE
M0 0.000 0.000 0.000 9.992
M1 0.924 0.853 0.852 3.847
Note. M1 includes Weight
ANOVA
Model Sum of Squares df Mean Square F p
M1 Regression 8433.838 1 8433.838 569.910 < .001
Residual 1450.258 98 14.799
Total 9884.096 99
Note. M1 includes Weight
Note. The intercept model is omitted, as no meaningful information can be shown.
Coefficients
Model Unstandardized Standard Error Standardized t p
M0 (Intercept) 168.345 0.999 168.480 < .001
M1 (Intercept) 125.194 1.848 67.746 < .001
Weight 0.599 0.025 0.924 23.873 < .001
Descriptives
N Mean SD SE
Height 100 168.345 9.992 0.999
Weight 100 72.048 15.411 1.541

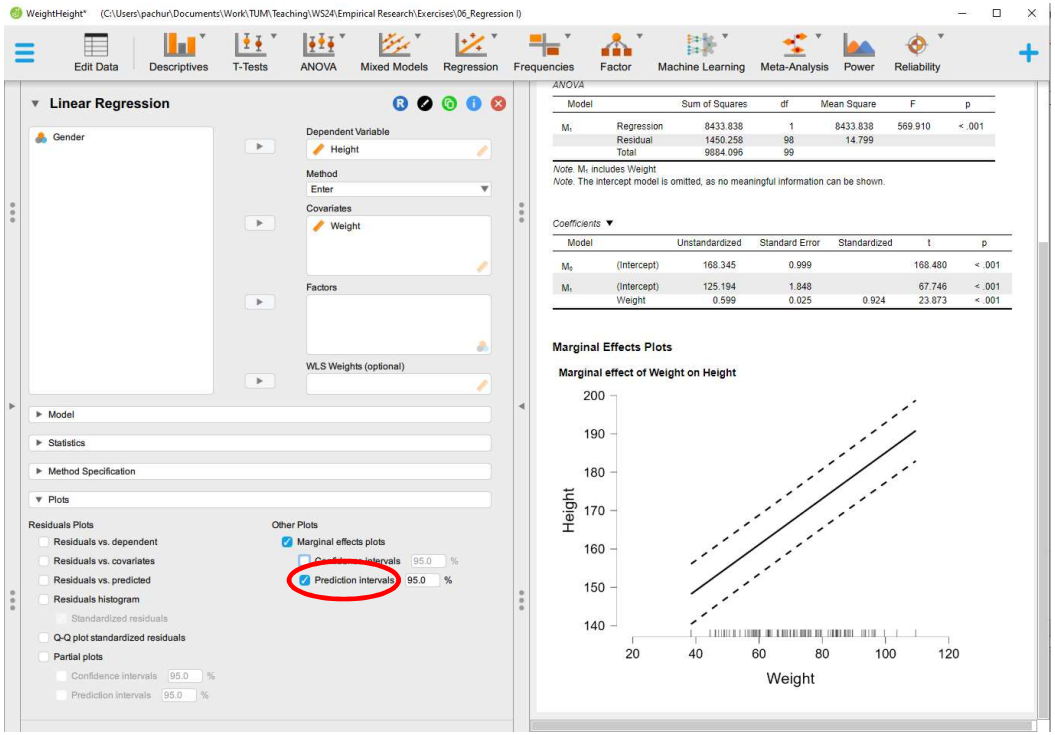

WeightHeight* (C:\Users\pachur\Documents\Work\TUM\Teaching\WS25\Empirical Research\Exercises\06_Regression I)
Edit Data
Descriptives
T-Tests
ANOVA
Mixed Models
Regression
Frequencies
Factor
Machine Learning
Meta-Analysis
Power
Process
Reliability
Factors
WLS Weights (optional)
Model
Statistics
Model Summary
R squared change
F change
AIC and BIC
Durbin-Watson
Coefficients
Estimates
From 5000 bootstraps
Confidence intervals 95 %
Tolerance and VIF
Vovk-Sellke maximum p-ratio
Display
Model fit
Descriptives
Part and partial correlations
Coefficients covariance matrix
Collinearity diagnostics
Residuals
Statistics
Casewise diagnostics
Std. residual > 3
Cook's dist. > 1
All
DFBETAS
DFFITS
Cov ratio
Leverage
Mahalanobis
Method Specification
Plots
Residuals Plots
Residuals vs. dependent
Residuals vs. covariates
Residuals vs. predicted
Residuals histogram
Standardized residuals
Q-Q plot standardized residuals
Partial plots
Confidence intervals 95 %
Prediction intervals 95 %
Other Plots
Marginal effects plots
Confidence intervals 95 %
Prediction intervals 95 %
Export
Weight 100 72.05 15.411 1.541
Residuals vs. Covariates
Residuals vs. Weight
Residuals
Standardized Residuals
Weight
Standardized Residuals Histogram
Density
Standardized Residuals
Q-Q Plot Standardized Residuals
Observed quantiles
Theoretical quantiles


WeightHeight* (C:\Users\pachur\Documents\Work\TUM\Teaching\WS24\Empirical Research\Exercises\06_Regression I)
Analyses
Synchronisation
Resize Data
Insert
Remove
Undo
Redo
Insert row above
Insert row below
Insert column after
Insert constructor column after
Insert R column before
Insert R column after
Name: Centered weight
Long name: Centered weight
Column type: Scale
Description:
Computed type: Computed with drag-and-drop
Computed column definition
Missing values
Label editor
+ - * ÷ / ^ √ % = ≠ < ≤ > ≥ ∧ ∨ | ¬
Gender
Height
Weight
Centered weight
Weight -...
Weight -mean(Weight)
Converting types
Compute column
Gender Height Weight Centered weight
1 Male 187.5714232 109.5777841 37.52937943
2 Male 174.7060363 73.52664405 1.478239377
3 Male 188.2396677 96.37160757 24.3232029
4 Male 182.1966851 99.67923905 27.63083438
5 Male 177.4997615 93.47645968 21.42805501
6 Male 170.8226598 68.95210656 -3.096298113
7 Male 174.7141064 83.31933354 11.27092887
8 Male 173.6052294 76.09091305 4.042508377
9 Male 170.2281321 79.6960365 7.647681827


WeightHeight* (C:\Users\pachur\Documents\Work\TUM\Teaching\WS24\Empirical Research\Exercises\06_Regression I)
Edit Data
Descriptives
T-Tests
ANOVA
Mixed Models
Regression
Frequencies
Factor
Machine Learning
Meta-Analysis
Power
Reliability
Centered weight
Factors
WLS Weights (optional)
Model
Statistics
Model Summary
R squared change
F change
AIC and BIC
Durbin-Watson
Coefficients
Estimates
From 5000 bootstraps
Confidence intervals 95.0 %
Tolerance and VIF
Vovk-Sellke maximum p-ratio
Display
Model fit
Descriptives
Part and partial correlations
Coefficients covariance matrix
Collinearity diagnostics
Residuals
Statistics
Casewise diagnostics
Std. residual > 3
Cook's dist. > 1
All
DFBETAS
DFFITS
Cov ratio
Leverage
Mahalanobis
Append residuals to data
Column name e.g., residuals
Method Specification
Plots
Results
Linear Regression
Model Summary - Height
Model R R² Adjusted R² RMSE
M0 0.000 0.000 0.000 9.992
M1 0.924 0.853 0.852 3.847
Note. M1 includes Centered weight
ANOVA
Model Sum of Squares df Mean Square F p
M1 Regression 8433.838 1 8433.838 569.910 < .001
Residual 1450.258 98 14.799
Total 9884.096 99
Note. M1 includes Centered weight
Note. The intercept model is omitted, as no meaningful information can be shown.
Coefficients
Model Unstandardized Standard Error Standardized t p
M0 (Intercept) 168.345 0.999 168.480 < .001
M1 (Intercept) 168.345 0.385 437.613 < .001
Centered weight 0.599 0.025 0.924 23.873 < .001
Descriptives
N Mean SD SE
Height 100 168.345 9.992 0.999
Centered weight 100 −7.673×10−17 15.411 1.541

Confidence and prediction intervals

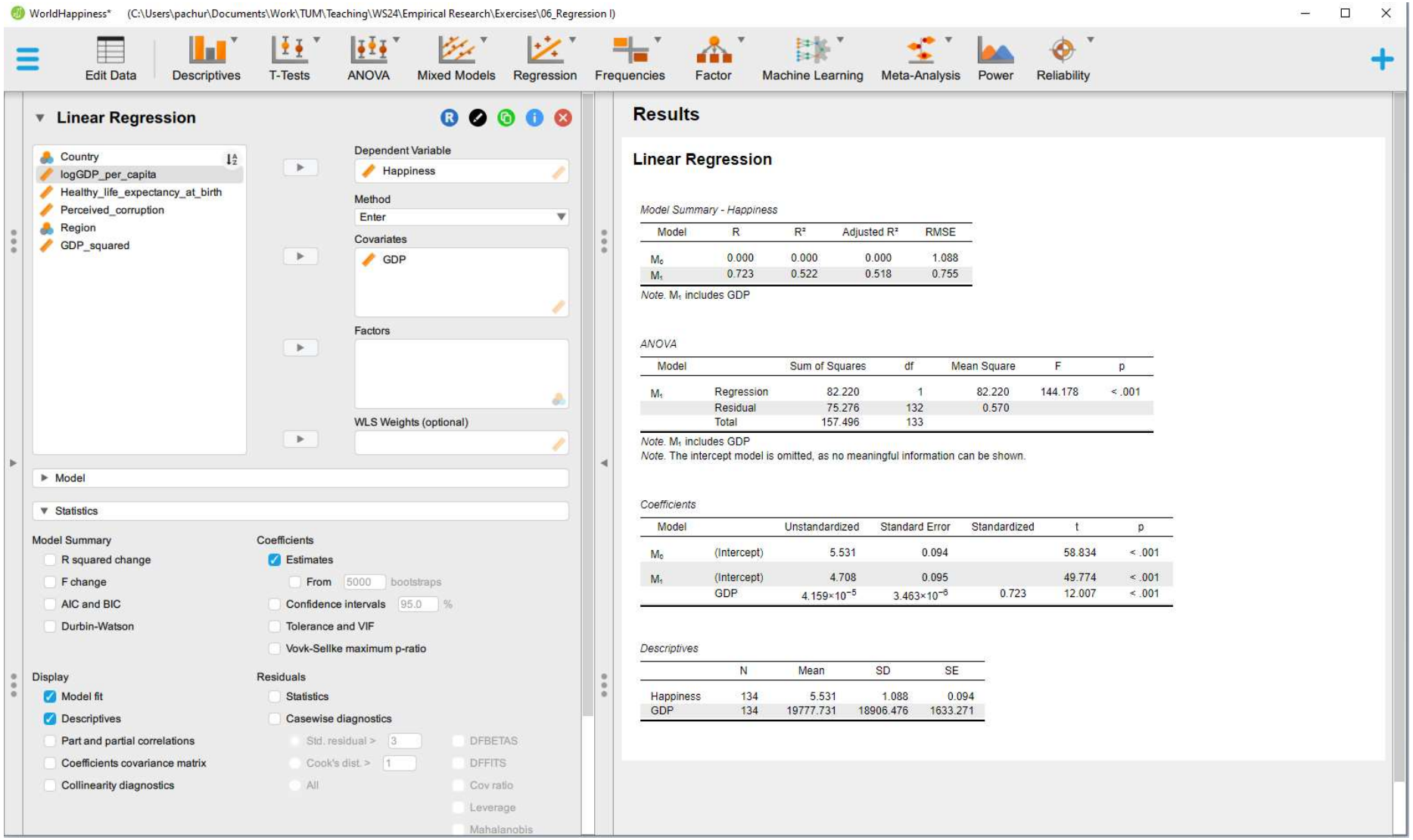


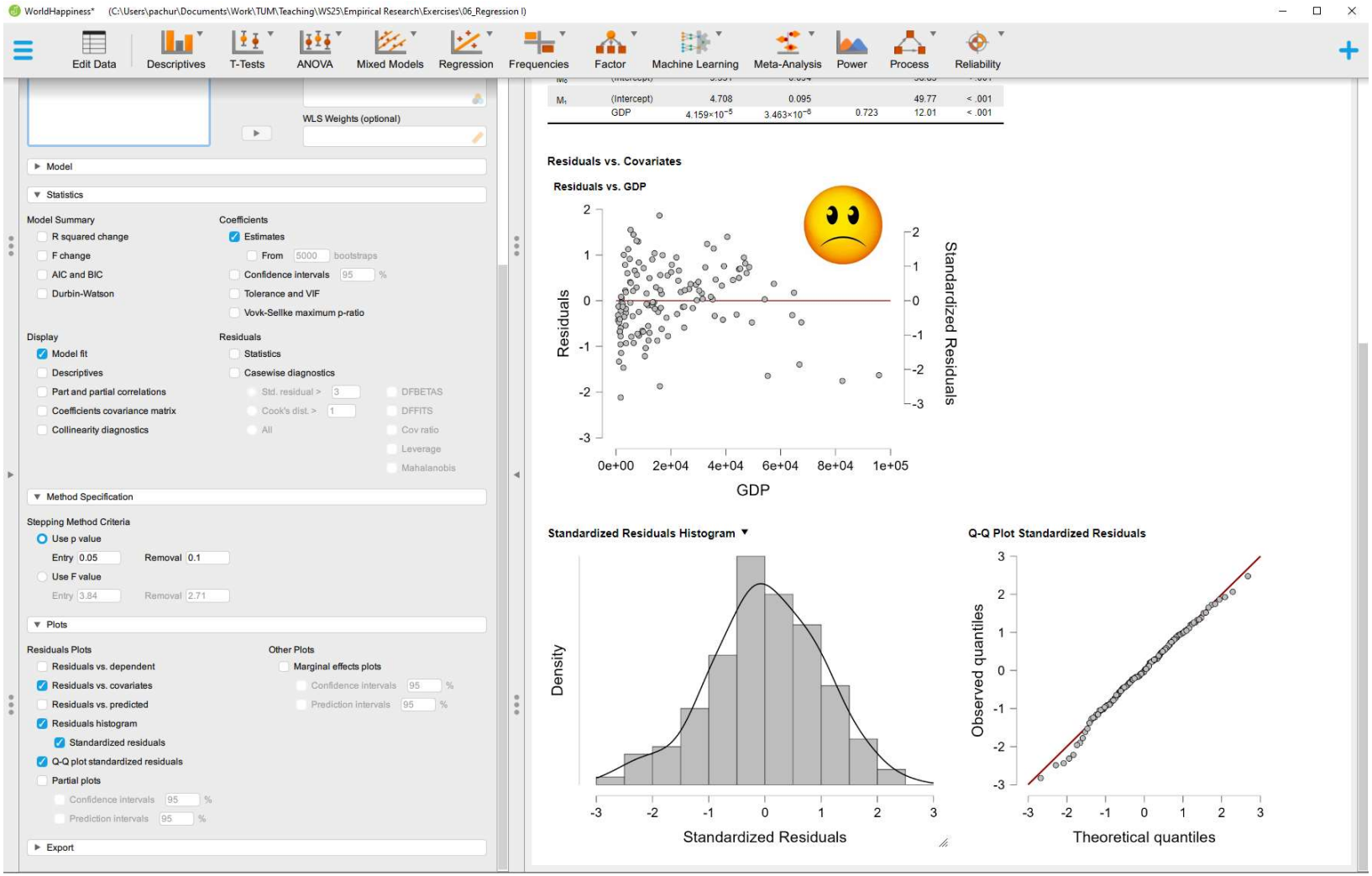


Exercise II

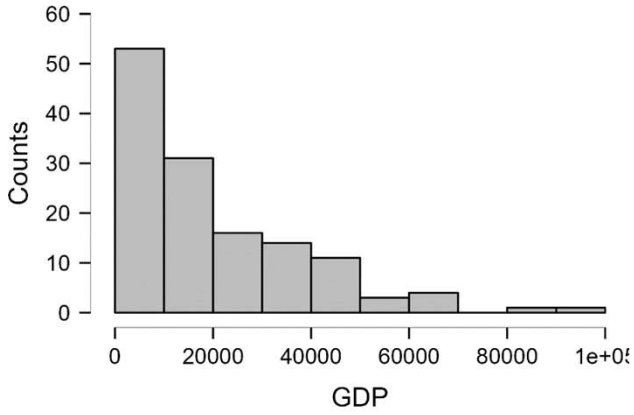
- Open data file “WorldHappiness.csv“
- Conduct regression analysis: Predict happiness from GDP
  - Regression coefficient, significance test, R^2
  - Test assumptions: linearity, homoscedasticity, and normality
- Conduct regression analysis: Predict happiness from logGDP
  - Regression coefficient, significance test, R^2
  - Test assumptions: linearity, homoscedasticity, and normality

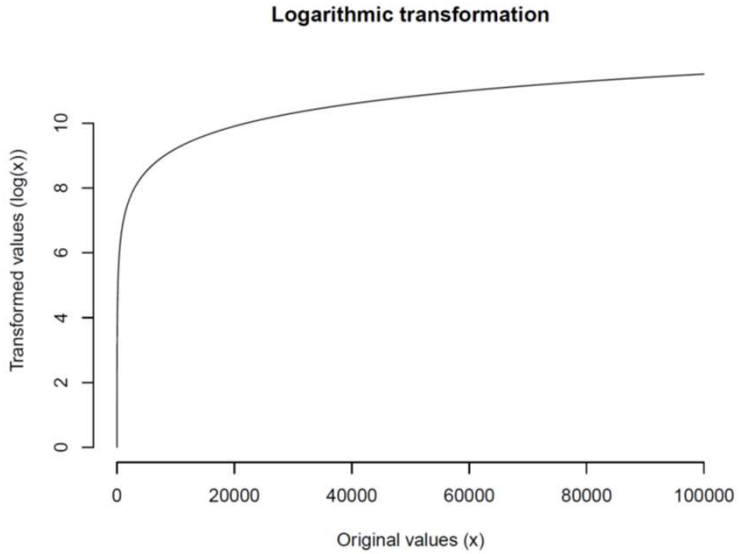
→ Compare the results

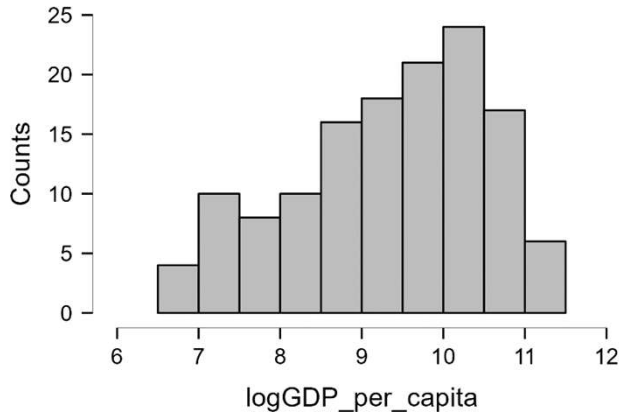

WorldHappiness* (C:\Users\pachur\Documents\Work\TUM\Teaching\WS24\Empirical Research\Exercises\06_Regression I)
Edit Data
Descriptives
T-Tests
ANOVA
Mixed Models
Regression
Frequencies
Factor
Machine Learning
Meta-Analysis
Power
Reliability
Linear Regression
Country
logGDP_per_capita
Healthy_life_expectancy_at_birth
Perceived_corruption
Region
GDP_squared
Dependent Variable
Happiness
Method
Enter
Covariates
GDP
Factors
WLS Weights (optional)
Model
Statistics
Model Summary
R squared change
F change
AIC and BIC
Durbin-Watson
Coefficients
Estimates
From 5000 bootstraps
Confidence intervals 95.0 %
Tolerance and VIF
Vovk-Sellke maximum p-ratio
Display
Model fit
Descriptives
Part and partial correlations
Coefficients covariance matrix
Collinearity diagnostics
Residuals
Statistics
Casewise diagnostics
Std. residual > 3
Cook's dist. > 1
All
DFBETAS
DFFITS
Cov ratio
Leverage
Mahalanobis
Results
Linear Regression
Model Summary - Happiness
Model R R² Adjusted R² RMSE
M0 0.000 0.000 0.000 1.088
M1 0.723 0.522 0.518 0.755
Note. M1 includes GDP
ANOVA
Model Sum of Squares df Mean Square F p
M1 Regression 82.220 1 82.220 144.178 < .001
Residual 75.276 132 0.570
Total 157.496 133
Note. M1 includes GDP
Note. The intercept model is omitted, as no meaningful information can be shown.
Coefficients
Model Unstandardized Standard Error Standardized t p
M0 (Intercept) 5.531 0.094 58.834 < .001
M1 (Intercept) 4.708 0.095 49.774 < .001
GDP 4.159×10−5 3.463×10−6 0.723 12.007 < .001
Descriptives
N Mean SD SE
Happiness 134 5.531 1.088 0.094
GDP 134 19777.731 18906.476 1633.271

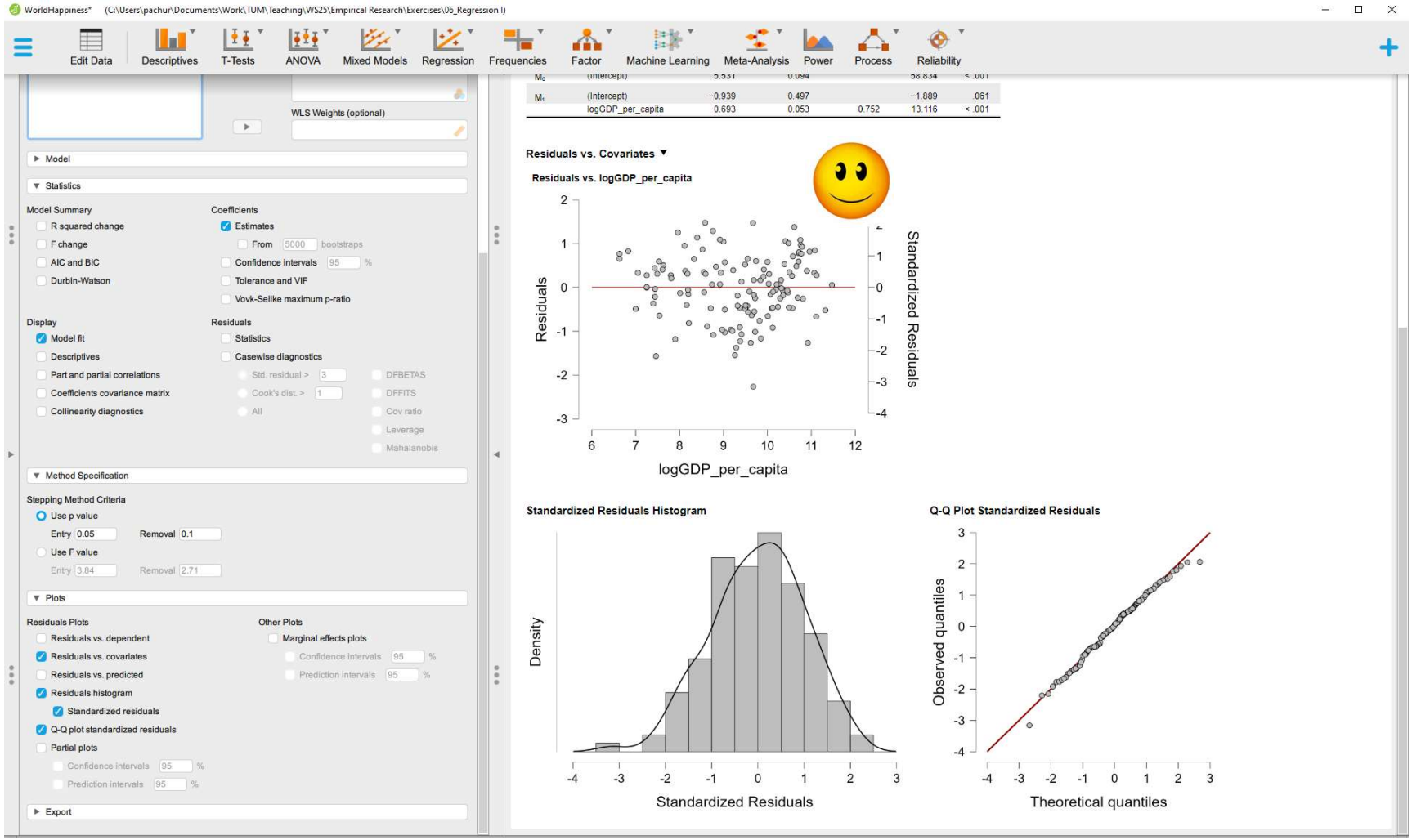

WorldHappiness* (C:\Users\pachur\Documents\Work\TUM\Teaching\WS25\Empirical Research\Exercises\06_Regression I)
Edit Data
Descriptives
T-Tests
ANOVA
Mixed Models
Regression
Frequencies
Factor
Machine Learning
Meta-Analysis
Power
Process
Reliability
WLS Weights (optional)
Model
Statistics
Model Summary
R squared change
F change
AIC and BIC
Durbin-Watson
Coefficients
Estimates
From 5000 bootstraps
Confidence intervals 95 %
Tolerance and VIF
Vovk-Sellke maximum p-ratio
Display
Model fit
Descriptives
Part and partial correlations
Coefficients covariance matrix
Collinearity diagnostics
Residuals
Statistics
Casewise diagnostics
Std. residual > 3
Cook's dist. > 1
All
DFBETAS
DFFITS
Cov ratio
Leverage
Mahalanobis
Method Specification
Stepping Method Criteria
Use p value
Entry 0.05
Removal 0.1
Use F value
Entry 3.84
Removal 2.71
Plots
Residuals Plots
Residuals vs. dependent
Residuals vs. covariates
Residuals vs. predicted
Residuals histogram
Standardized residuals
Q-Q plot standardized residuals
Partial plots
Confidence intervals 95 %
Prediction intervals 95 %
Other Plots
Marginal effects plots
Confidence intervals 95 %
Prediction intervals 95 %
Export
M1 (Intercept) 4.708 0.095 49.77 < .001
GDP 4.159×10−5 3.463×10−6 0.723 12.01 < .001
Residuals vs. Covariates
Residuals vs. GDP
Residuals
Standardized Residuals
GDP
0e+00 2e+04 4e+04 6e+04 8e+04 1e+05
Standardized Residuals Histogram
Density
Standardized Residuals
Q-Q Plot Standardized Residuals
Observed quantiles
Theoretical quantiles

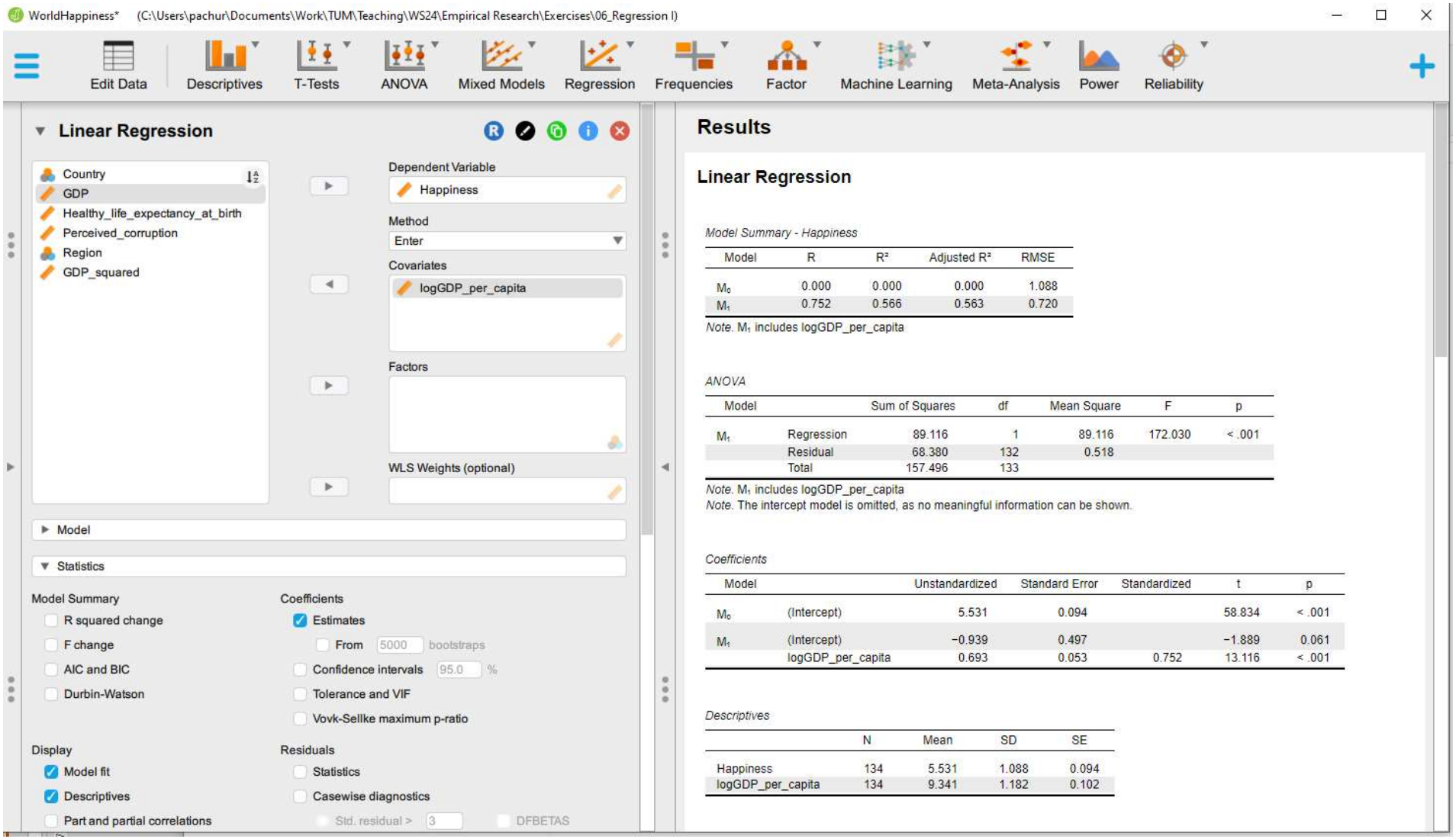
Logarithmic transformation








WorldHappiness* (C:\Users\pachur\Documents\Work\TUM\Teaching\WS25\Empirical Research\Exercises\06_Regression I)
Edit Data
Descriptives
T-Tests
ANOVA
Mixed Models
Regression
Frequencies
Factor
Machine Learning
Meta-Analysis
Power
Process
Reliability
WLS Weights (optional)
Model
Statistics
Model Summary
R squared change
F change
AIC and BIC
Durbin-Watson
Coefficients
Estimates
From 5000 bootstraps
Confidence intervals 95 %
Tolerance and VIF
Vovk-Sellke maximum p-ratio
Display
Model fit
Descriptives
Part and partial correlations
Coefficients covariance matrix
Collinearity diagnostics
Residuals
Statistics
Casewise diagnostics
Std. residual > 3
Cook's dist. > 1
All
DFBETAS
DFFITS
Cov ratio
Leverage
Mahalanobis
Method Specification
Stepping Method Criteria
Use p value
Entry 0.05
Removal 0.1
Use F value
Entry 3.84
Removal 2.71
Plots
Residuals Plots
Residuals vs. dependent
Residuals vs. covariates
Residuals vs. predicted
Residuals histogram
Standardized residuals
Q-Q plot standardized residuals
Partial plots
Confidence intervals 95 %
Prediction intervals 95 %
Other Plots
Marginal effects plots
Confidence intervals 95 %
Prediction intervals 95 %
Export
(Intercept) 5.531 0.094 58.834 < .001
M1 (Intercept) −0.939 0.497 −1.889 .061
logGDP_per_capita 0.693 0.053 0.752 13.116 < .001
Residuals vs. Covariates
Residuals vs. logGDP_per_capita
Residuals
Standardized Residuals
logGDP_per_capita
Standardized Residuals Histogram
Density
Standardized Residuals
Q-Q Plot Standardized Residuals
Observed quantiles
Theoretical quantiles


WorldHappiness* (C:\Users\pachur\Documents\Work\TUM\Teaching\WS24\Empirical Research\Exercises\06_Regression I)
Edit Data
Descriptives
T-Tests
ANOVA
Mixed Models
Regression
Frequencies
Factor
Machine Learning
Meta-Analysis
Power
Reliability
Linear Regression
Country
GDP
Healthy_life_expectancy_at_birth
Perceived_corruption
Region
GDP_squared
Dependent Variable
Happiness
Method
Enter
Covariates
logGDP_per_capita
Factors
WLS Weights (optional)
Model
Statistics
Model Summary
R squared change
F change
AIC and BIC
Durbin-Watson
Coefficients
Estimates
From 5000 bootstraps
Confidence intervals 95.0 %
Tolerance and VIF
Vovk-Sellke maximum p-ratio
Display
Model fit
Descriptives
Part and partial correlations
Residuals
Statistics
Casewise diagnostics
Std. residual > 3
DFBETAS
Results
Linear Regression
Model Summary - Happiness
Model R R² Adjusted R² RMSE
M0 0.000 0.000 0.000 1.088
M1 0.752 0.566 0.563 0.720
Note. M1 includes logGDP_per_capita
ANOVA
Model Sum of Squares df Mean Square F p
M1 Regression 89.116 1 89.116 172.030 < .001
Residual 68.380 132 0.518
Total 157.496 133
Note. M1 includes logGDP_per_capita
Note. The intercept model is omitted, as no meaningful information can be shown.
Coefficients
Model Unstandardized Standard Error Standardized t p
M0 (Intercept) 5.531 0.094 58.834 < .001
M1 (Intercept) −0.939 0.497 −1.889 0.061
logGDP_per_capita 0.693 0.053 0.752 13.116 < .001
Descriptives
N Mean SD SE
Happiness 134 5.531 1.088 0.094
logGDP_per_capita 134 9.341 1.182 0.102

Exercise III

- A study has collected data on the earnings (in $) and height (in cm) of N = 1140 female employees ($SD_{earning}$ = $15,508, SD_{height} = 6.54 cm). A regression analysis predicting earnings from height yields an unstandardized regression weight for height of b = 148.57 (p = .034).
- Use G*Power to conduct an a priori power analysis to determine the required sample size for a study aiming to replicate the observed effect, with α = .05 and a power of .80.

A priori power analysis for simple regression