statgraphics®
centurion

Quantile Regression

Presented by
Dr. Neil W. Polhemus

Quantile Regression

- Constructs linear models for predicting specified quantiles.

- Useful when:
  - primary interest concerns a percentile of the distribution rather than the mean.
  - the distribution of the data at a specified combination of the predictor variables is not Gaussian.
  - the variance of Y depends on X.
  - there are outliers present.

Applications

- Growth curves
- Ecology
- Epidemiology
- Health services utilization
- CEO pay
- Household income
- Home prices
- Sea ice extent
- Astrophysics
- Chemistry
- Genomics
- Waiting times
- Product reliability

statgraphics®
statgraphics.com centurion

Basic Model Structure

$$Q_\tau(Y) = \beta_0(\tau) + \beta_1(\tau)X_1 + \beta_2(\tau)X_2 + \cdots + \beta_p(\tau)X_p + \epsilon$$

$Q_\tau(Y)$: conditional τ–th quantile of dependent variable Y

$X_1, X_2, \ldots X_p$: predictor variables

Note that the coefficients depend on τ.

statgraphics®
statgraphics.com centurion

Statgraphics

- Statgraphics uses the *quantreg* program in R to fit models.

- *Quantreg* was written by R. Koenker.

- You should download the latest build (19.2.02) which includes a few tweaks to the *Quantile Regression* procedure.

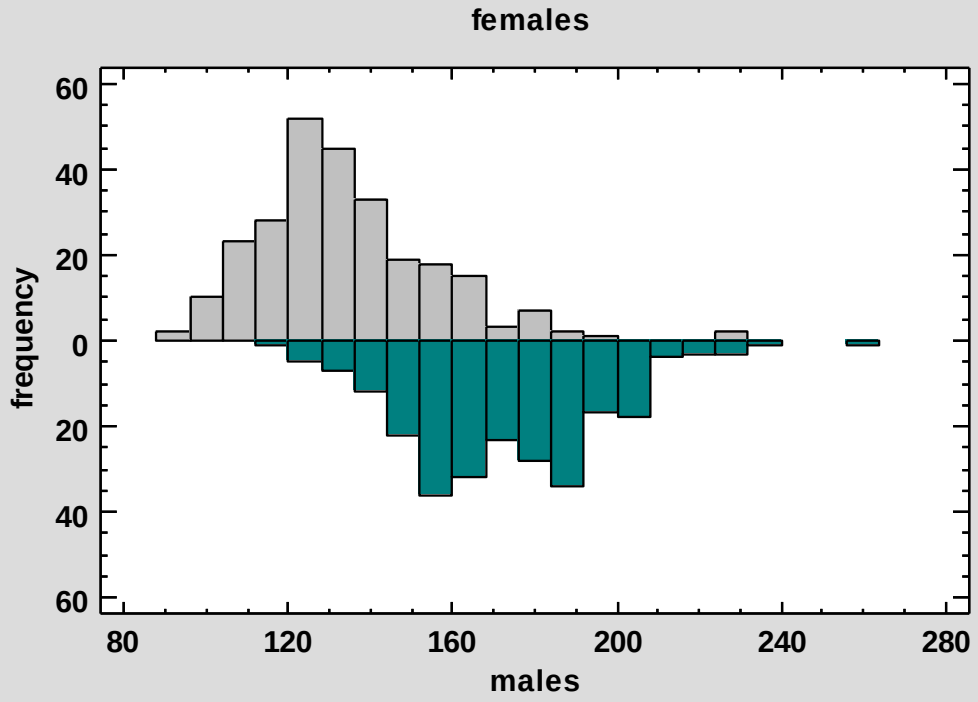
Example #1

- First example is taken from the Journal of Statistics Education Data Archive.

- Information about 247 men and 267 women sampled at fitness centers in California:
  - 21 body dimension measurements
  - age
  - height
  - weight
  - gender

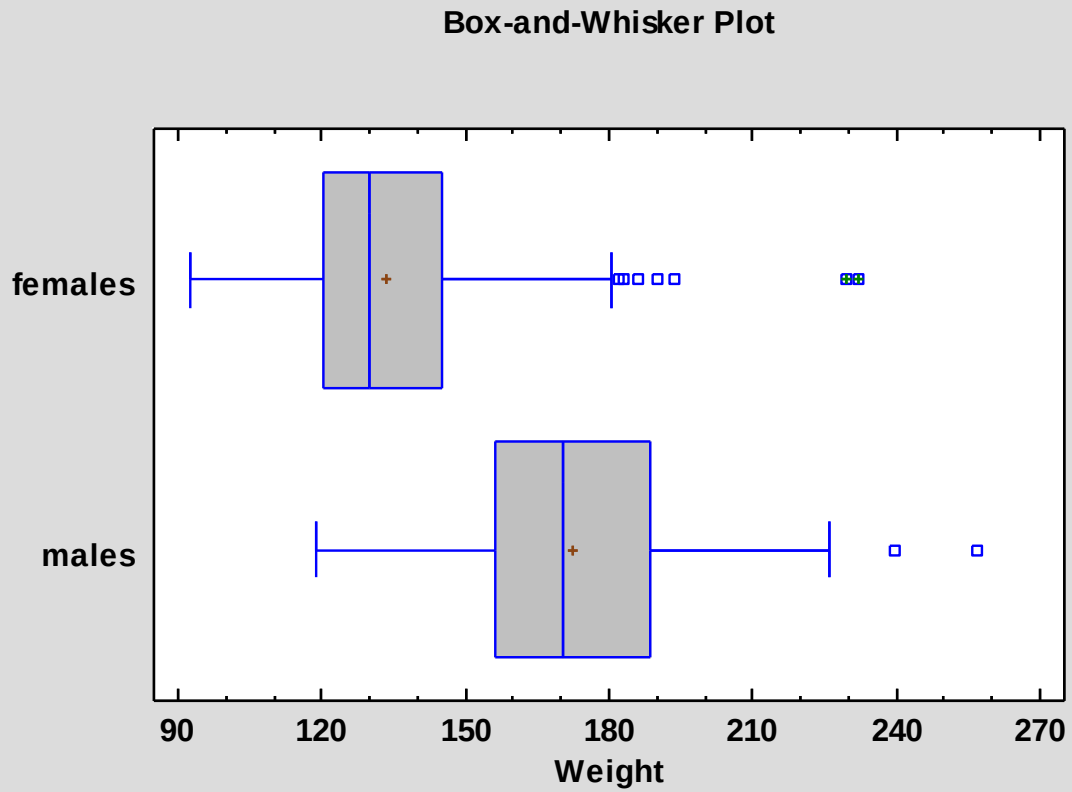
statgraphics®
statgraphics.com centurion

Histogram of Weight



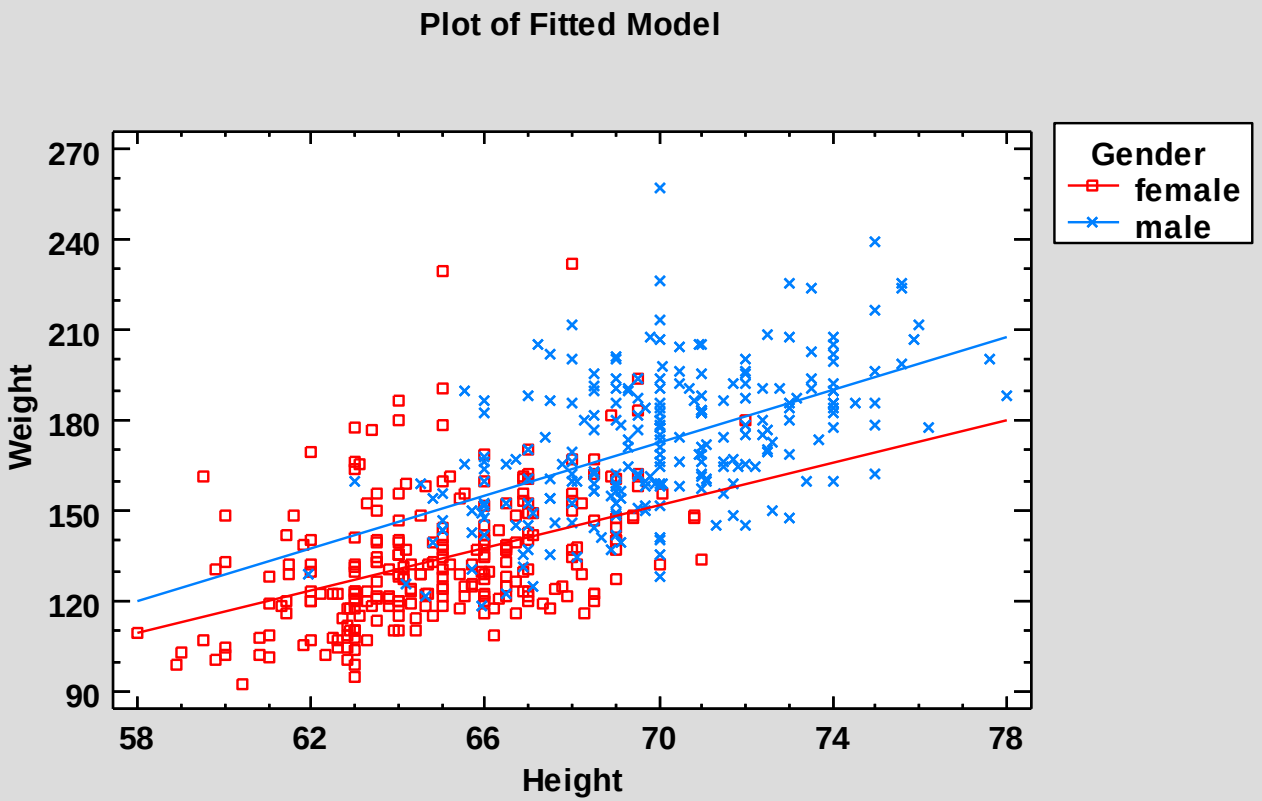
statgraphics®
statgraphics.com
centurion

Box and Whisker Plot of Weight



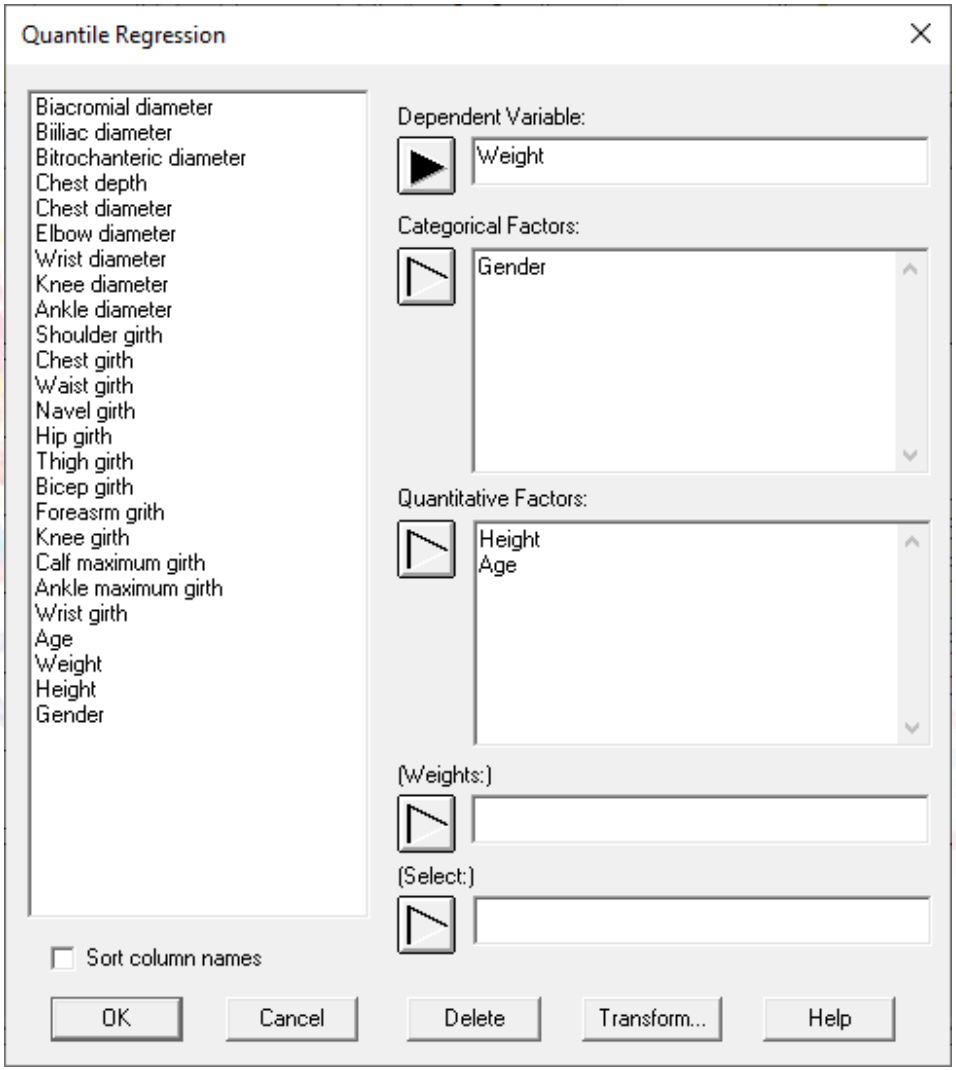
statgraphics®
statgraphics.com
centurion

Comparison of Regression Lines



statgraphics®
statgraphics.com centurion

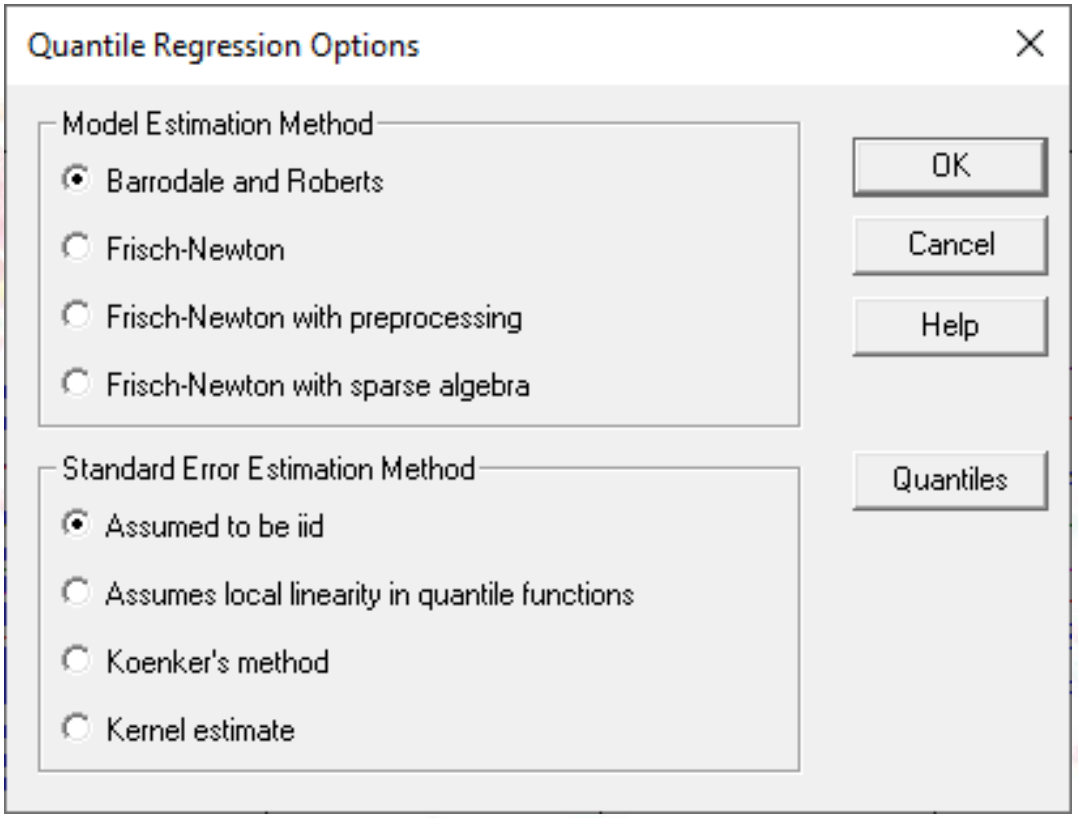
Data Input Dialog Box



statgraphics® centurion

statgraphics.com

Analysis Options Dialog Box



statgraphics®
statgraphics.com centurion

Methods

- *Barrodale and Roberts* – for up to several thousand observations.

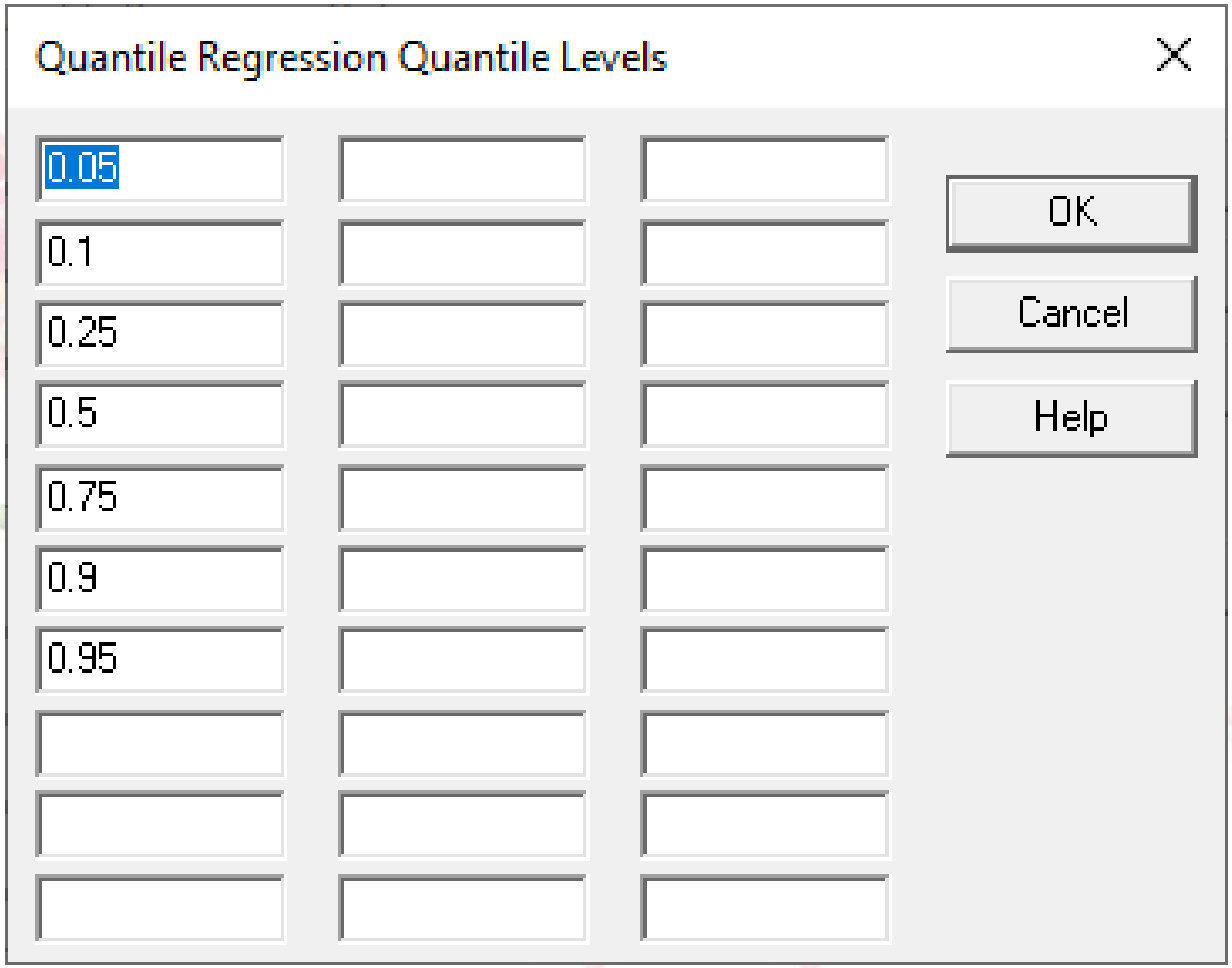
- *Frisch-Newton* – useful for larger problems.

- *Frisch-Newton with preprocessing* – useful for even larger problems. Best for large *n* and small *p*.

- *Frisch-Newton with sparse algebra* – useful for large *n* and large *p*.

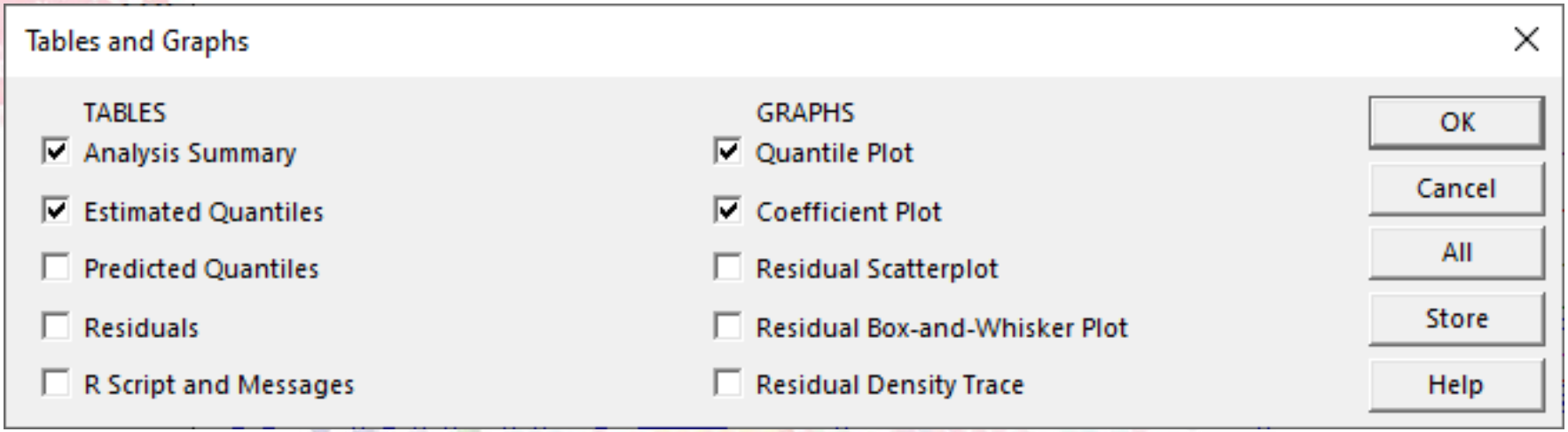
statgraphics®
statgraphics.com centurion

Quantile Dialog Box



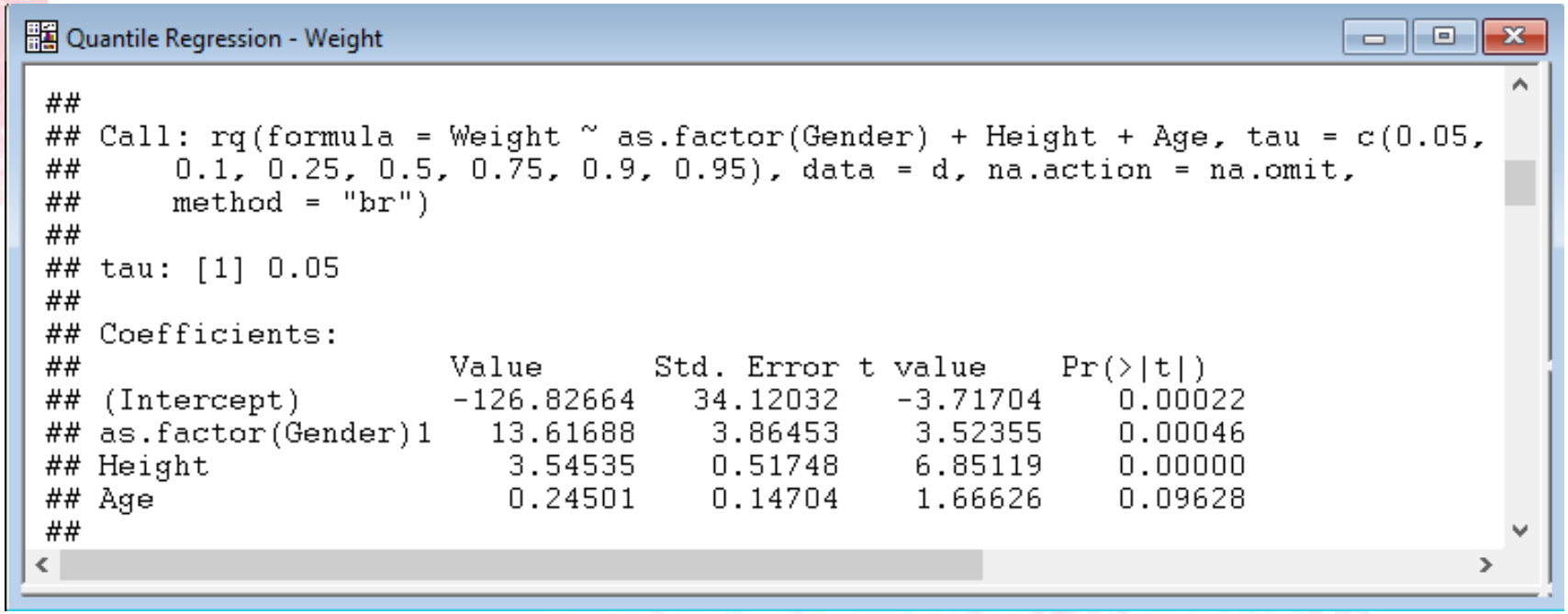
statgraphics®
statgraphics.com centurion

Tables and Graphs



statgraphics®
statgraphics.com centurion

Analysis Summary



statgraphics®
statgraphics.com centurion

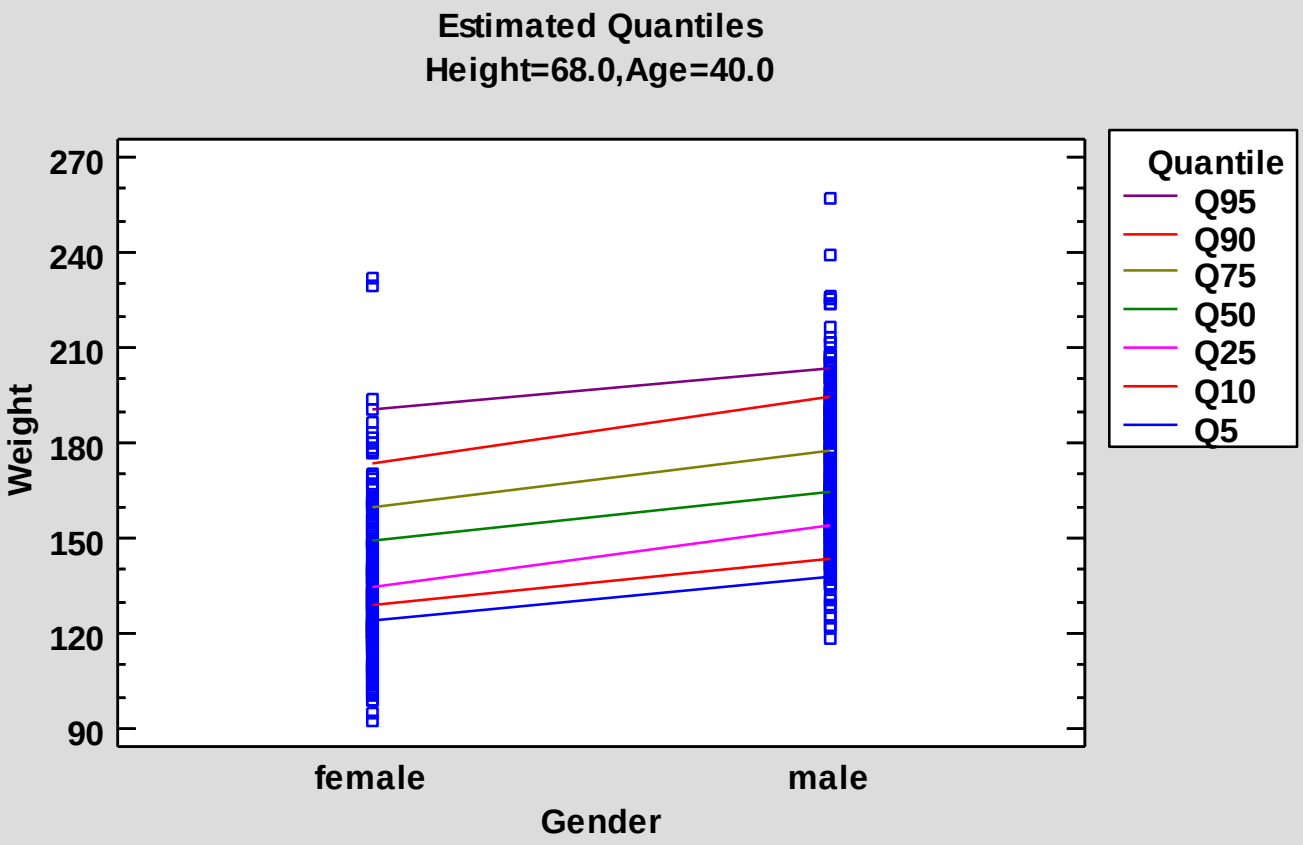
Analysis Summary

Quantile Regression - Weight

```
##
## Call: rq(formula = Weight ~ as.factor(Gender) + Height + Age, tau = c(0.05,
##     0.1, 0.25, 0.5, 0.75, 0.9, 0.95), data = d, na.action = na.omit,
##     method = "br")
##
## tau: [1] 0.5
##
## Coefficients:
##                     Value      Std. Error t value    Pr(>|t|)
## (Intercept)        -169.75731   23.59772   -7.19380    0.00000
## as.factor(Gender)1   15.73834    2.67272    5.88851    0.00000
## Height                4.45059    0.35789   12.43564    0.00000
## Age                   0.39921    0.10169    3.92564    0.00010
##
```

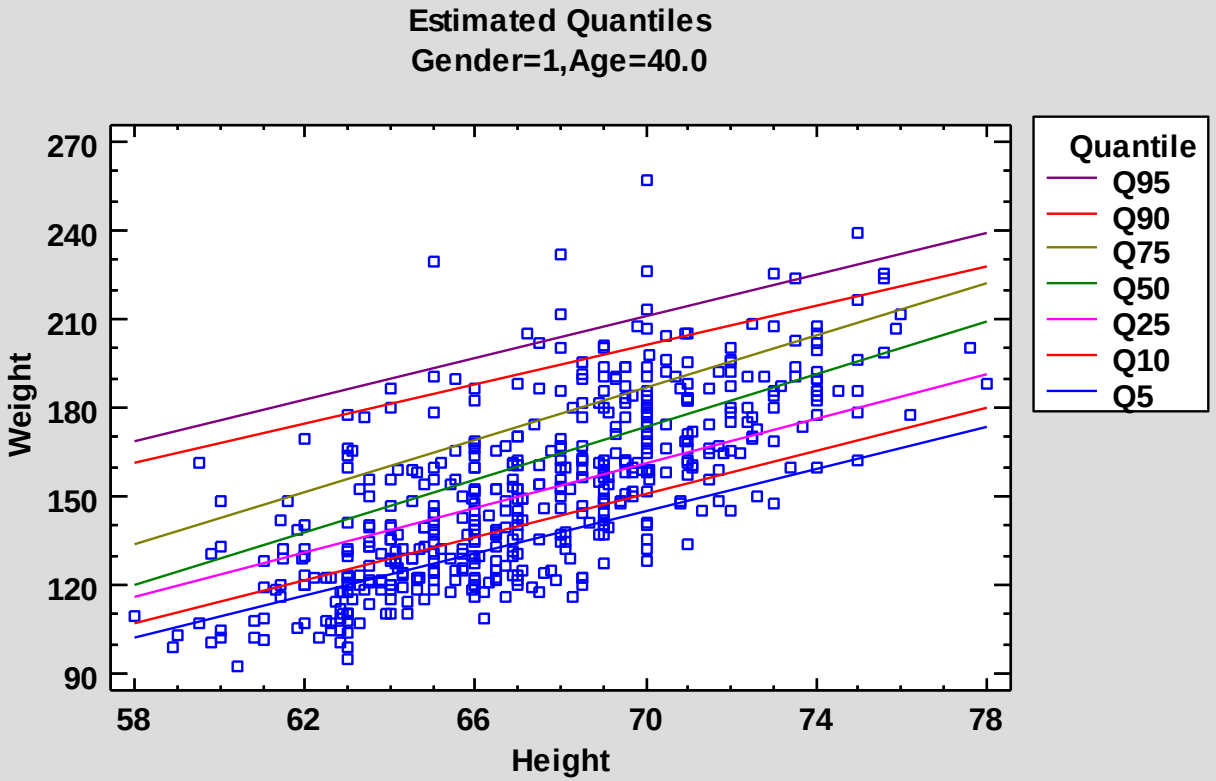
statgraphics® centurion
statgraphics.com

Effect of Gender



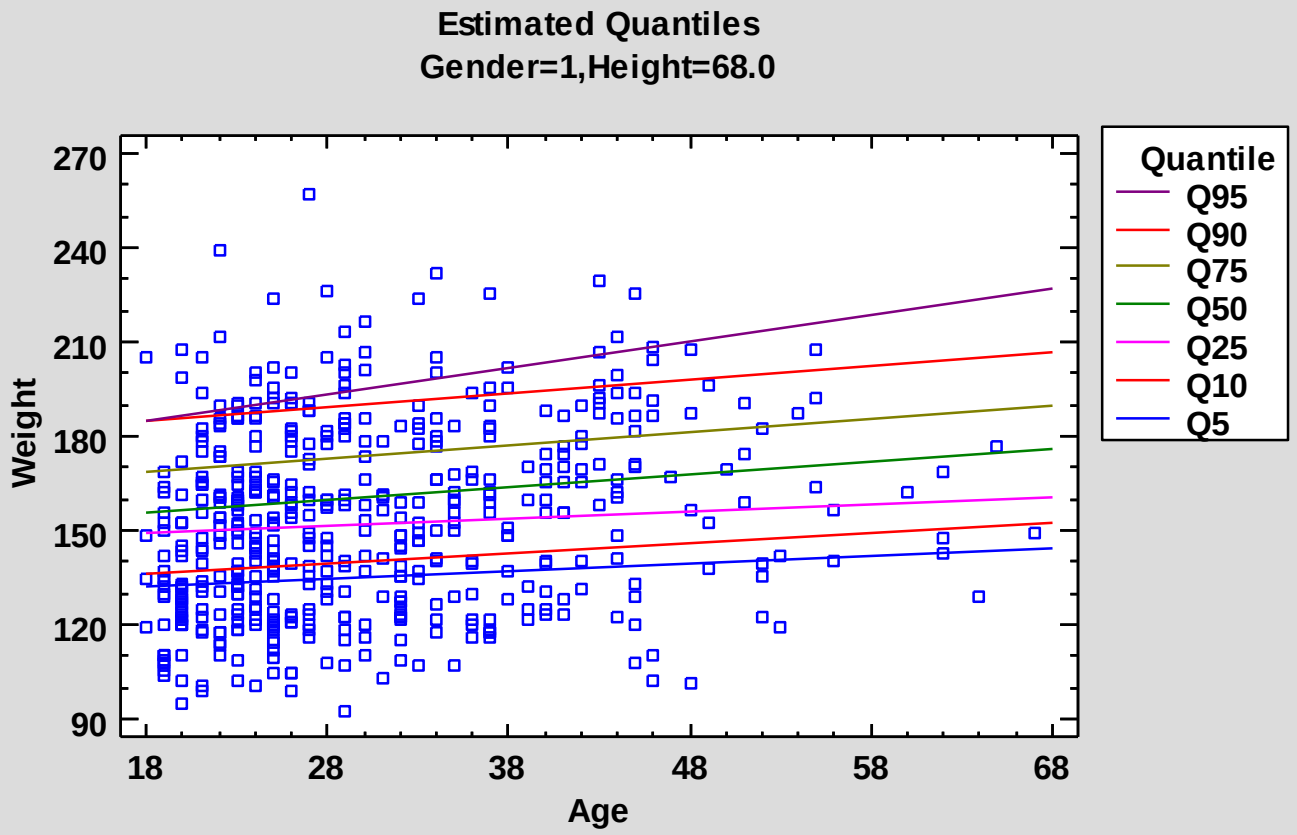
statgraphics®
statgraphics.com centurion

Effect of Height



statgraphics®
statgraphics.com centurion

Effect of Age



statgraphics®
statgraphics.com centurion

Estimated Quantiles

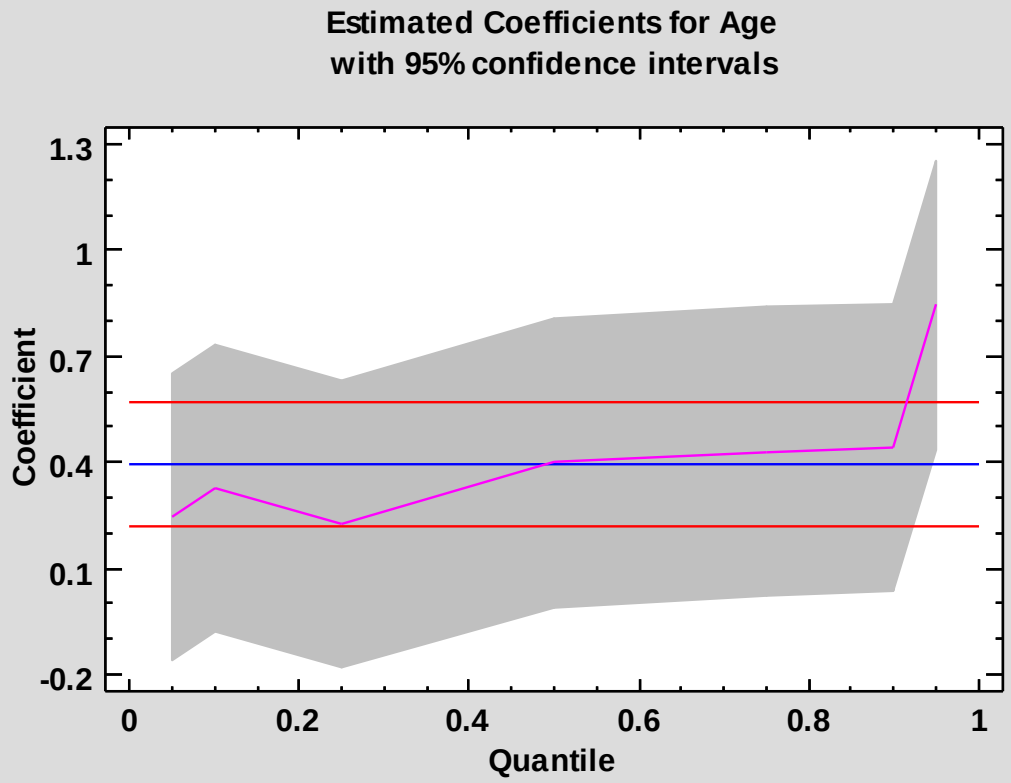
Quantile Regression - Weight

Estimated Quantiles

Row	Weight	Gender	Height	Age	Q5	Q10	Q25	Q50	Q75	Q90	Q95
1	144.6	1.0	68.5	21.0	134.792	139.228	151.483	159.23	171.951	187.948	189.403
2	158.3	1.0	69.0	23.0	137.055	141.721	153.809	162.254	175.023	190.5	192.862
3	177.9	1.0	76.2	28.0	163.806	169.811	181.966	196.294	209.008	216.707	222.57
4	160.1	1.0	73.4	23.0	152.654	157.884	170.33	181.836	194.475	205.163	208.436
5	173.8	1.0	73.7	22.0	153.473	158.658	171.232	182.772	195.371	205.72	208.654
6	164.9	1.0	71.5	21.0	145.428	150.248	162.747	172.582	185.214	197.946	200.022
7	190.5	1.0	72.4	26.0	149.844	155.196	167.249	178.583	191.347	203.159	207.43
8	172.9	1.0	72.6	27.0	150.798	156.258	168.225	179.873	192.662	204.268	208.983
9	136.7	1.0	68.9	23.0	136.7	141.354	153.434	161.809	174.581	190.167	192.508
10	179.9	1.0	72.4	21.0	148.619	153.554	166.126	176.587	189.192	200.945	203.208
11	168.9	1.0	70.9	23.0	143.791	148.701	160.943	170.71	183.423	196.832	199.587
12	184.3	1.0	70.0	22.0	140.355	145.067	157.34	166.305	179.013	193.39	195.557
13	198.5	1.0	75.6	20.0	159.719	164.981	177.917	190.43	202.909	211.167	213.69
14	164.5	1.0	69.3	26.0	138.853	143.808	155.609	164.787	177.642	192.828	196.457
15	156.6	1.0	68.5	23.0	135.282	139.885	151.932	160.028	172.813	188.834	191.092
16	175.5	1.0	72.4	22.0	148.864	153.883	166.351	176.987	189.623	201.388	204.052
17	206.8	1.0	75.9	30.0	163.232	169.365	181.289	195.757	208.544	216.593	223.197
18	154.4	1.0	67.5	22.0	131.492	135.883	147.953	155.179	167.961	185.059	186.708
19	159.6	1.0	68.1	29.0	135.334	140.385	151.777	160.643	173.63	190.156	194.743
20	189.4	1.0	69.3	22.0	137.873	142.495	154.711	163.19	175.918	191.057	193.079
21	173.8	1.0	69.3	22.0	137.873	142.495	154.711	163.19	175.918	191.057	193.079

statgraphics®
statgraphics.com centurion

Comparing Quantile Coefficients



statgraphics®
statgraphics.com centurion

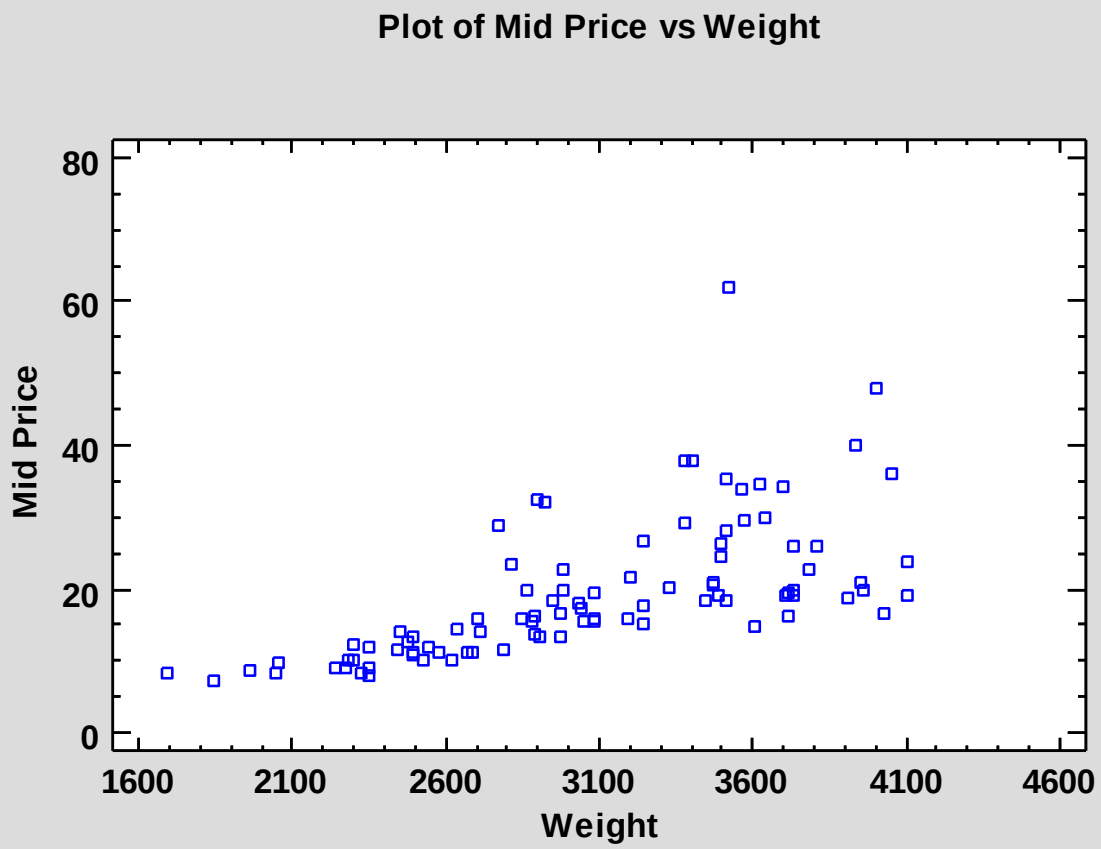
Example #2

- Second example is also taken from the Journal of Statistics Education Data Archive.

- Information about 93 makes and models of automobiles manufactured in 1993
  - Price of automobile
  - Weight

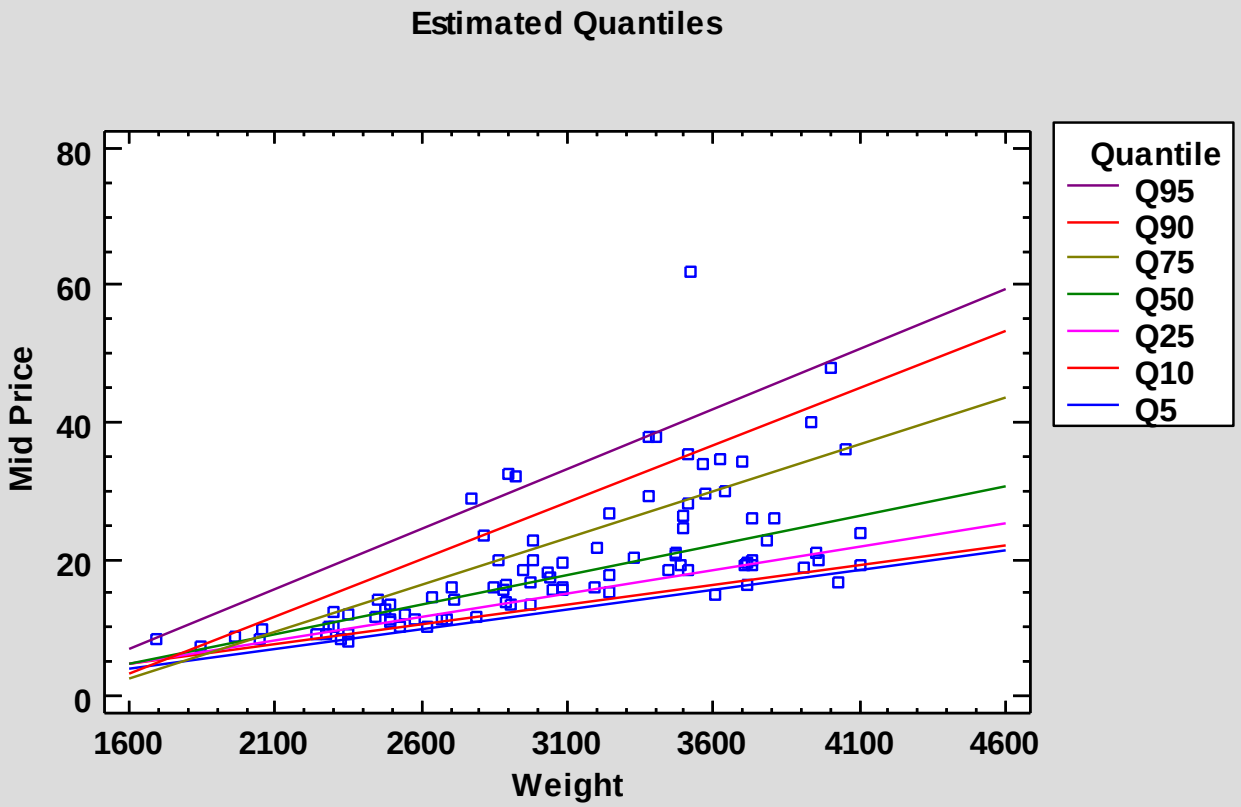
statgraphics®
statgraphics.com centurion

Scatterplot



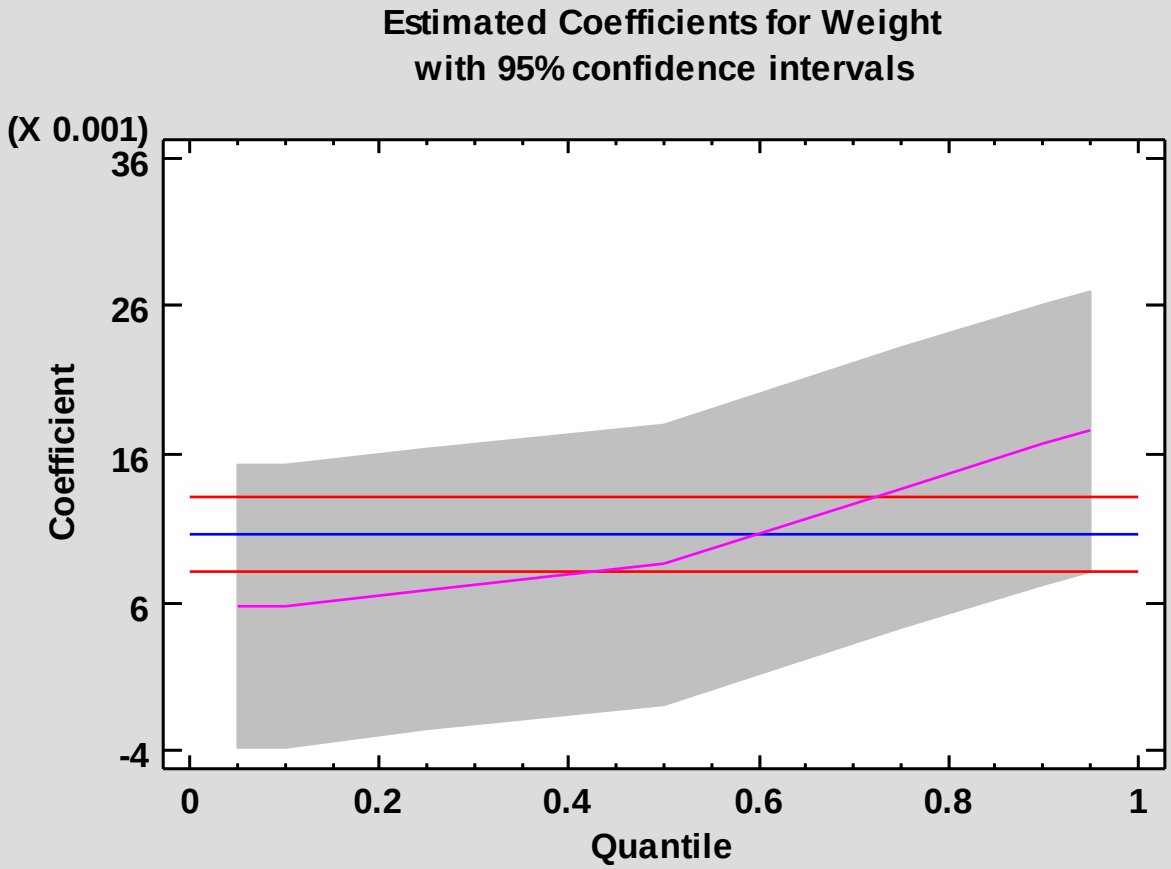
statgraphics®
statgraphics.com centurion

Estimated Quantiles



statgraphics®
statgraphics.com centurion

Comparing Quantile Coefficients



statgraphics®
statgraphics.com centurion

References

- StatFolios and data files are at: *www.statgraphics.com/webinars*
- R Package "quantreg" (2021) https://cran.r-project.org/web/packages/quantreg/quantreg.pdf
- Quantile Regression (2005) Roger Koenker. (Econometric Society Monographs, Series Number 38)
- Body and car data from: http://jse.amstat.org/jse_data_archive.htm