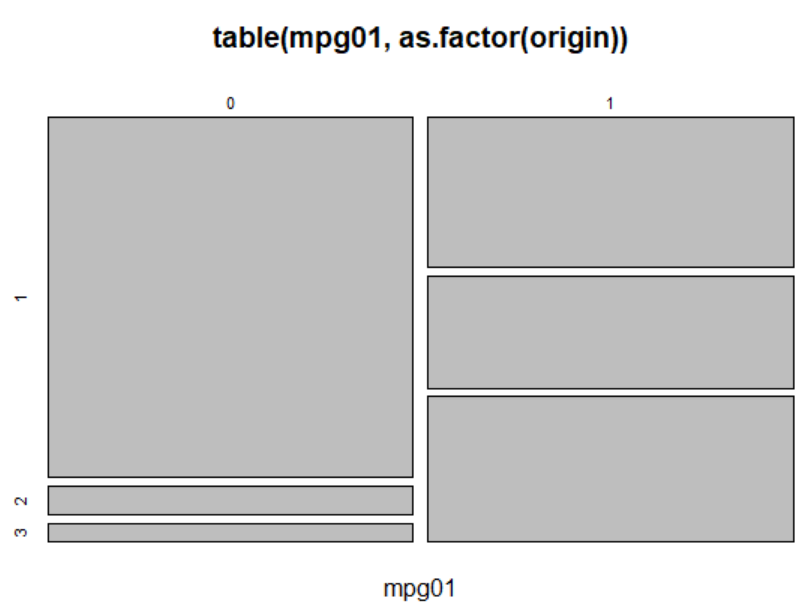
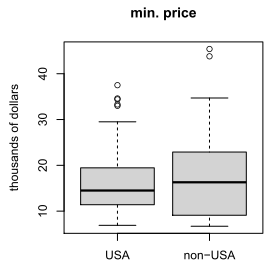
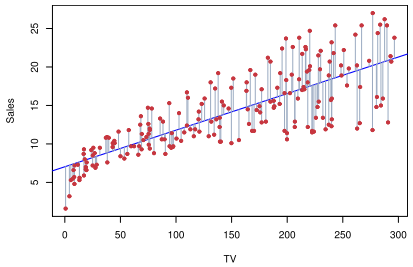
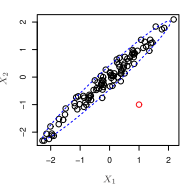
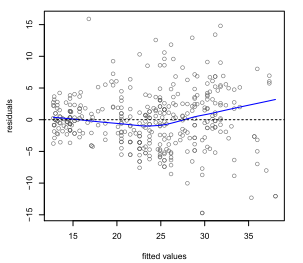
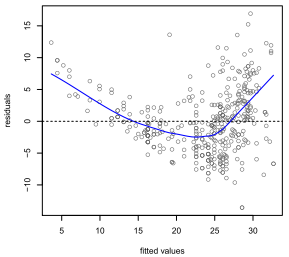
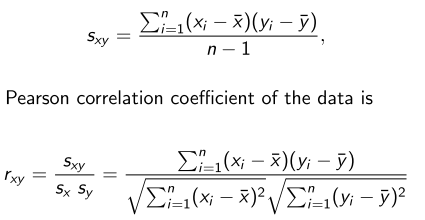
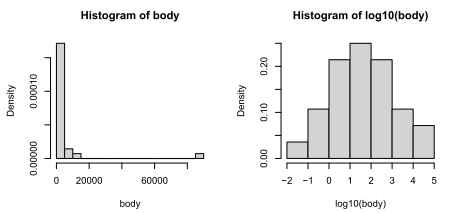
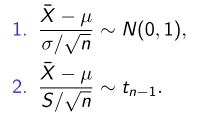
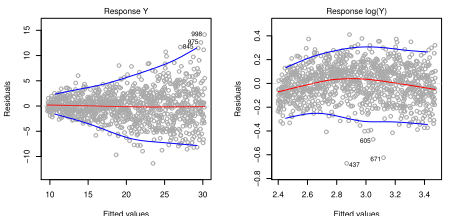
1. Plots & Tables:
   1. Pie chart
   2. Bar chart
   3. Scatterplots (with smoothing)
   4. Mosaicplot  
      
   5. QQ plot to check normality of the distributions (slide 113 / 267)
   6. Boxplots  
      
   7. Histograms & estimation of density
      1. (Bimodal distribution!)
      2. Skewness (negative, symm, positive)
   8. Linear model:
      1. Results of model  
         
      2. QQ plot for analysis of residuals (to check assumptions correctness for linear model)
      3. High leverage points plot (68 / 98)  
         
      4. Residual plot (58 / 98 part II) (**Residual – diff between y\_i and estimated y^\_i. Unlike errors, residuals do not have same variance. If residuals are dispersed around horizontal axis – linear regression is appropriate**). Check for funnel shape (66 / 98). For example, non-linearity is indicated from residual plot below. 🡪 in this case try non-linearity:, or transforming Y variable with log, sqrt 
2. Tables
   1. Table representation of some data
   2. Covariance matrix
   3. Covariation matrix
   4. Correlation matrix (to check collinearity (70 / 98))
3. Numeric characterizations:
   1. IQR, quantiles
   2. Standard deviation
   3. Covariance of two variables & correlation:  
      
   4. Linear models:
      1. RSE, residual standard error (slide 21 / 98 part II)
      2. R^2 statistics (always increased when new predictors are added)
      3. Adjusted R^2 (54 / 98 part II)  
         
      4. Standard error of predictors’ estimate (42 / 98 part II)
      5. Leverage statistic (69 / 98)
4. Transformations:
   1. Log  
      
   2. Root (square, cubical)
   3. Box & Cox transformation
   4. Inverse (1 / x) transformation
5. Statistical tests & hypothesis
   1. Confidence interval mean estimation (160 / 267)
   2. Hypotheses for inference of variance (197 / 267)
   3. T-test (paired / non-paired)
   4. Test for equal variances
   5. Hypothesis for mean testing (179 / 267)  
      
6. Our models
   1. Simple linear regression (**Assumed, that errors are normal(mean = 0, variance = sigma^2) for multiple linear regression**) (examples – 39 / 98 part II)
   2. Multilinear model selection:
      1. Forward selection (adding predictors to empty model)
      2. Backward selection (removing from full model)
      3. Mixed selection
   3. Linear model with non-linear relation between columns (62 / 98) (**And check on the residual plot, if its better (64 / 98)**)
   4. Try non-linear transformations of the predictor (log(Y) for radioactive decay, sqrt(Y)) 
   5. Model with interaction effects
   6. ANOVA model (85 / 98)
   7. ANCOVA model