UNSCRAMBLER: MLR Lecture. (Updated for the X-version)

Part 1 (after slide 2)

1. Import data – we use the uploaded Unscrambler version that includes names
2. Mark Price: –>plot – >histogram
3. With/without statistics: Click Statistics icon in top menu bar
4. Potentially Normal probability plot
5. Make a LOGPRICE variable:
    a. Insert a blank column, say, as number 1.
    b. Copy the Price into this (and potentially mark it)
    c. Task->Transform-> Compute General: "V1=log(V1)" (Change name into LOGPRICE)
6. Select DATE and LOGPRICE: Mark one by clicking on variable number, mark the 2nd by "Ctrl" click.
7. Make 2D scatter plot: -> plot ->scatter plot
8. Make scatter-plots of LOGPRICE vs. DISTANCE and ELEVATION

Part 2 (after slide 11)

1. Compute correlation matrix: ->Task -> Analyze -> Descriptive Statistics
2. Make column ranges to prepare the analysis of : Y=LOGPRICE and X (The 7 other variables): -> Edit -> Define Ranges: Name a Range and choose the interval accordingly – one for Y and one for X. Click "Update" to save a range under the chosen name.
3. Do the MLR: ->Task->Analyze->Multiple Linear Regression: Chose the relevant X and Y, and click OK twice.
4. Look at ANOVA table: Top left corner.

Part 3 (after slide 12)

1. Look at the regression coefficients!
2. Remove Non-significant terms one by one: Go to regression coef. Plot, use "Mark one by one" to mark the most non-significant variable and then – rightclick the MLR object in the project navigator -> Recalculate without marked.
3. Interpret! (Check DISTANCE – compare with simple analysis)(DISTANCE is related to ELEVATION)

Part 4 (after slide 17/18)

1. Plot residuals: Residuals vs. predicted, normal probability plots.
2. Influence plot

Part 5 (after slide 20)

1. (Potentially: Look under "Validation". Copy Y-residuals column into the X-data set and plot residuals versus X'es)
2. Construct e.g. DISTANCE^2 and COUNTY*ELEVATION
3. Run the MLR including these effects.