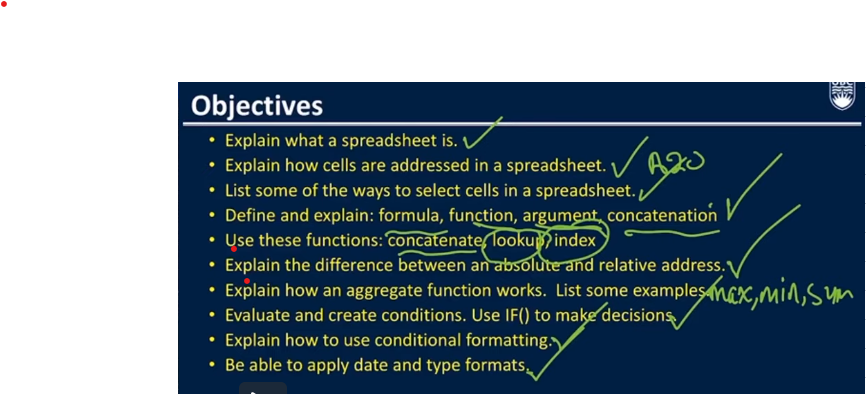
DATA 301 Notes September 16th

* Check instructions in syllabus for software
* Microsoft Excel
  + Good for quick data analysis, sharing
* Sorting data—select “Sort” option under Data menu
  + Highlight columns you wish to sort
* Filtering Data—shows a subset of the rows in the spreadsheet that passes a given condition
  + Select “auto filter” under the data then filter menu
* A cell has 2 components—an address (column, row) and location (that can store a number, text, or formula)
  + Referencing – column name and row number (e.g. D15)
  + Select cells—hold Shift and use arrow keys
  + Paste formulas
  + String concentration—2 or more strings are combined by appending them in order
    - Concatenate
* Check course content for more stuff
* How to multiply
  + $ = absolute address
  + E.g. if cell A1 has “=C5 + D4)”
    - To ensure the same formula is used in further cells, write “$C$5 + $D$4”
* Aggregate functions
  + Allow you to perform a summary function over a range of values (data values or cell locations)
  + E.g. (MIN, MAX, SUM, AVERAGE, COUNT, MEDIAN)
  + Colons indicate range (e.g. D2:D12=from D2 to D12)
* Column width adjusting
  + Can be done using click and drag, Format bar at top, or select all columns and double click
* Conditional Formatting
* Date and Type Formats
  + Check “more number formats”
* 
* Data cleaning
  + Data-
* Charts
  + Graphical representation of spreadsheet data
  + Select “insert”, then click “charts” icon, and pick the chart type
  + Chart tools—allows you to modify the data in the chart, change the chart type, and move the chart in the Worksheet
  + Trendlines--- easily added to chart
  + Sparklines—a tiny chart in a worksheet cell for a quick data overview – a chart within a single cell
  + Pivot tables—allow for easily aggregating and exploring large data sets
* Goal Seek—not on test
  + - Way to solve for a variable given the target value of another cell
* Analysis ToolPak—excel add-in that has a set of statistical and data analysis tools such as ANOVA, covariance, regression, and t-test
* Linear regression
  + Models the relationship between a dependant variable y and explanatory variables x
  + E.g. y= Bx + C
  + l