Now that the data is uploaded, run these commands to prepare it for making the charts:

```
library(tidyverse)

# input your file path #here
avgtuition <- read.csv(
  file = "...",
  stringsAsFactors = FALSE
)

head(avgtuition)
```

Navigate to Sheet 1 (bottom left corner).
Open your original csv file.
Tableau Public

Different points in time, highlight the data points in the

To change the colors, try dragging the data onto the colors under the Marks menu! Test it out, and see what happens. To adjust the color

Manually:

Excel is great for making quick and easy scatterplots! Let's make one in excel now!

There are different options for overcoming this issue, including carefully subsetting the data, utilizing transparency for the points (which we

Here is one of the visualizations he created: a smoothed correlation diagram of the heights of parents and children.

Denis suggest that the original scatterplot was made in 1833 by English scientist John Frederick W. Herschel; he created scatterplots to display

Scatterplots dominate scientific journals; Tufte estimated that more than 70% of all charts in scientific publications are scatterplots. Friendly and

The origin of the scatterplot is not entirely clear (Kopf 2018). They were historically called scatter diagrams and were often used to plot points on

Denis suggests that scatterplots are particularly useful for visualizing correlations among multiple variables, but if the points are also matched to a color, shape, or size, an additional variable can be represented in the visualization.

Scatterplots usually display values for two variables, but if the points are also matched to a color, shape, or size, an additional variable can be represented in the visualization. These graphs are primarily used for observing data and displaying relationships between at least two (generally numeric) variables. The points or

That it is hard to tell what the relationship is between the points and the variables. If the plot is extremely dense, then this might be occurring (Yi

Keep in mind…

11/11/2020

Digi Data Visualization: Lecture 3

6 rows

1.1.2.23.2.2019). There are different options. An example of correlation could be the relationship between square feet of a building and the use of electricity.

These graphs are primarily used for observing data and displaying relationships between at least two (generally numeric) variables. The points or

```r
#this can also be applied to the different variable values

ggplot(data = avgtuition, aes(x=Decade, y=Average.Tuition)) + geom_point(aes(size =

#adjust opacity of the points

#add in the main title

e F command above. Try using T to turn them back on and see what happens!

## Loading required package: carData

## Warning in library(carData, check = FALSE): package or data file not found

## Error in library(carData, check = FALSE): package or data file not found

#input your file path #here

```
Excel

We will learn how to make area graphs in Excel, Tableau, and R, also via zoom 4pm est on Nov. 18th to conclude our lecture series this semester.

Looking ahead!

Here is an example image of the output in Tableau, with the trend line!

\[ \text{computed for sum of Gr Liv Area given sum of Sale Price. The model may be significant at } p \leq 0.05. \]

Note: If you wish to add a "trend line", go to the top menu Analysis. Then select Trend Lines -> Show Trend Lines. If you click into this, you can.

\[ \text{Excel} \]

\[ \text{Tableau} \]

\[ \text{R} \]

## `geom_smooth()` using formula 'y ~ x'

els = scales::comma)+scale_y_continuous(labels = scales::comma)

```
ggplot(data = aptsales, aes(x=SalePrice, y=GrLivArea)) + geom_point(aes(color = SalePrice)) + labs(x=
```