# Week1 Asking The "Right Questions"

## 1.1 Introduction

### Tips for Becoming a Data Analyst

1, Ask Questions, Nourish Curiosity, and Embrace the Unknown

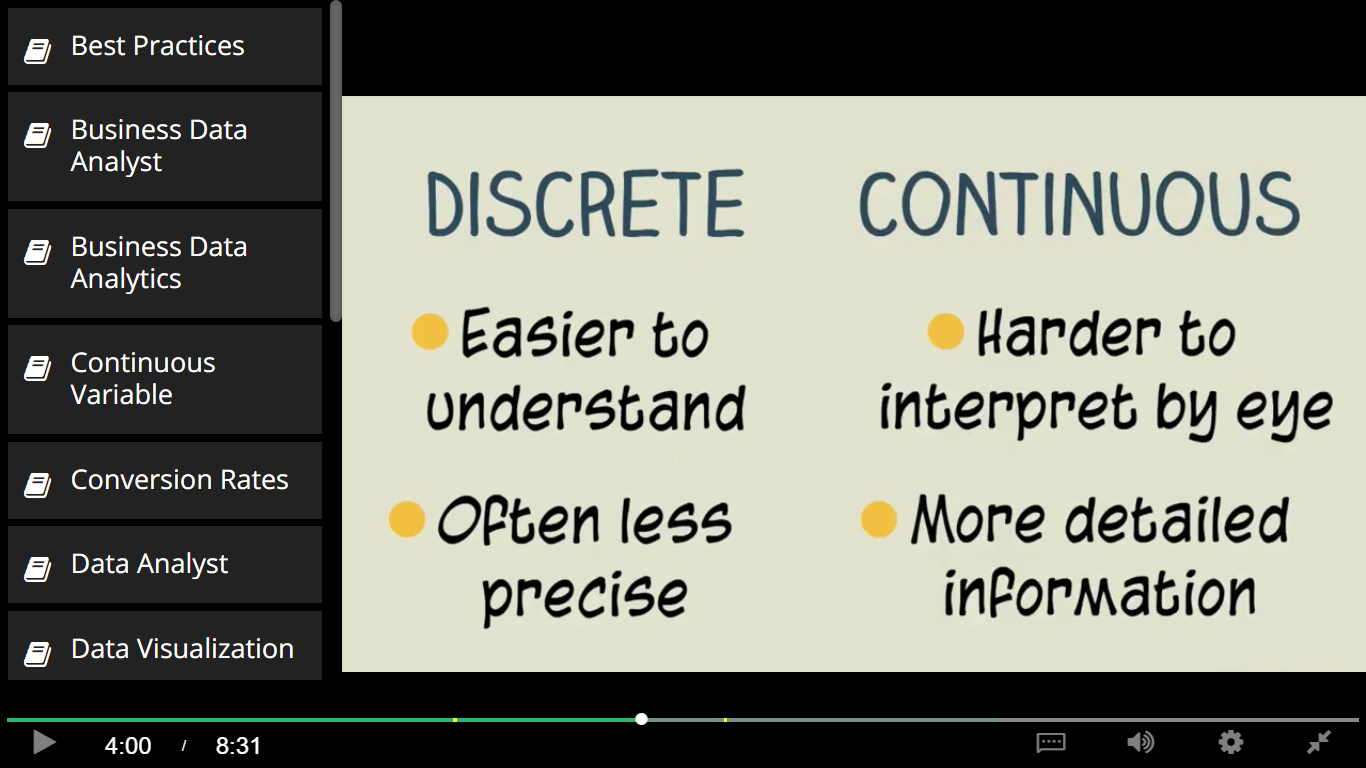
to stay competitive, your skill set will need to be dynamic

2, Start Thinking about Everything You See as a Dependent or Independent Variable.

the dependent variable, because it is the phenomenon I am measuring and I'm most interested in

3, explore the advantages and disadvantages of continuous versus discrete variables

Bar graph vs. line graph



4, make sure you always listen, and contribute

5. Train your skepticism muscles

whenever somebody tells you something with extreme confidence, temper your expectations

 6, that's why you need to seek details

 7, Cherish Precision

8, which is the the best practices in data analytics are not necessarily the most common practices in data analytics

9. the expectations of your teammates and stakeholders matter

10. the ability to put yourself in other people's shoes

## 1.2 Asking the Right Questions

### S.M.A.R.T. Objectives

Figure out where all your stakeholders’ true motivations and financial interests lie. Remember, unfortunately, they may not always know where their motivations and financial interest lie. Even, if they think they know.

What problem is this business having that you are hoping to solve by developing this project?

Can you tell me more about how this problem is affecting the business?

What is your ideal outcome for this project?

Specific

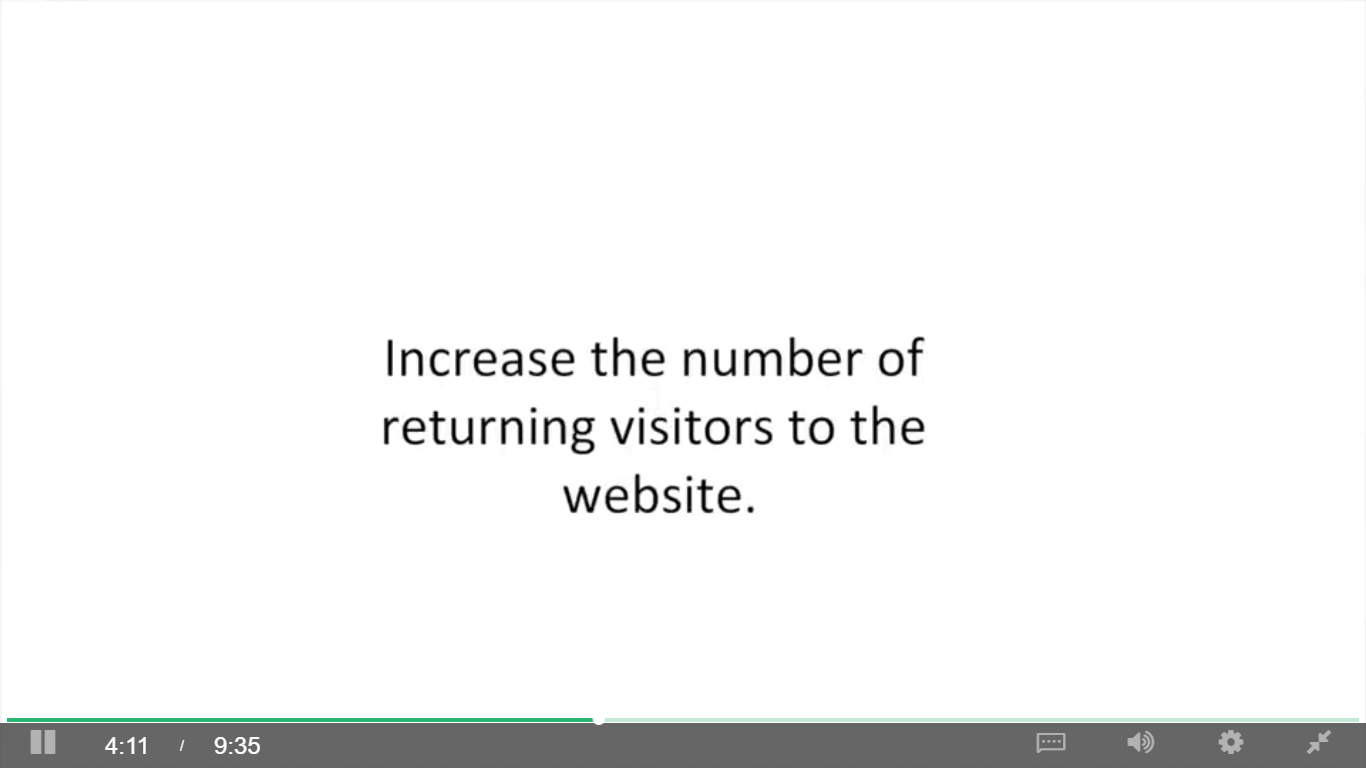
Measurable

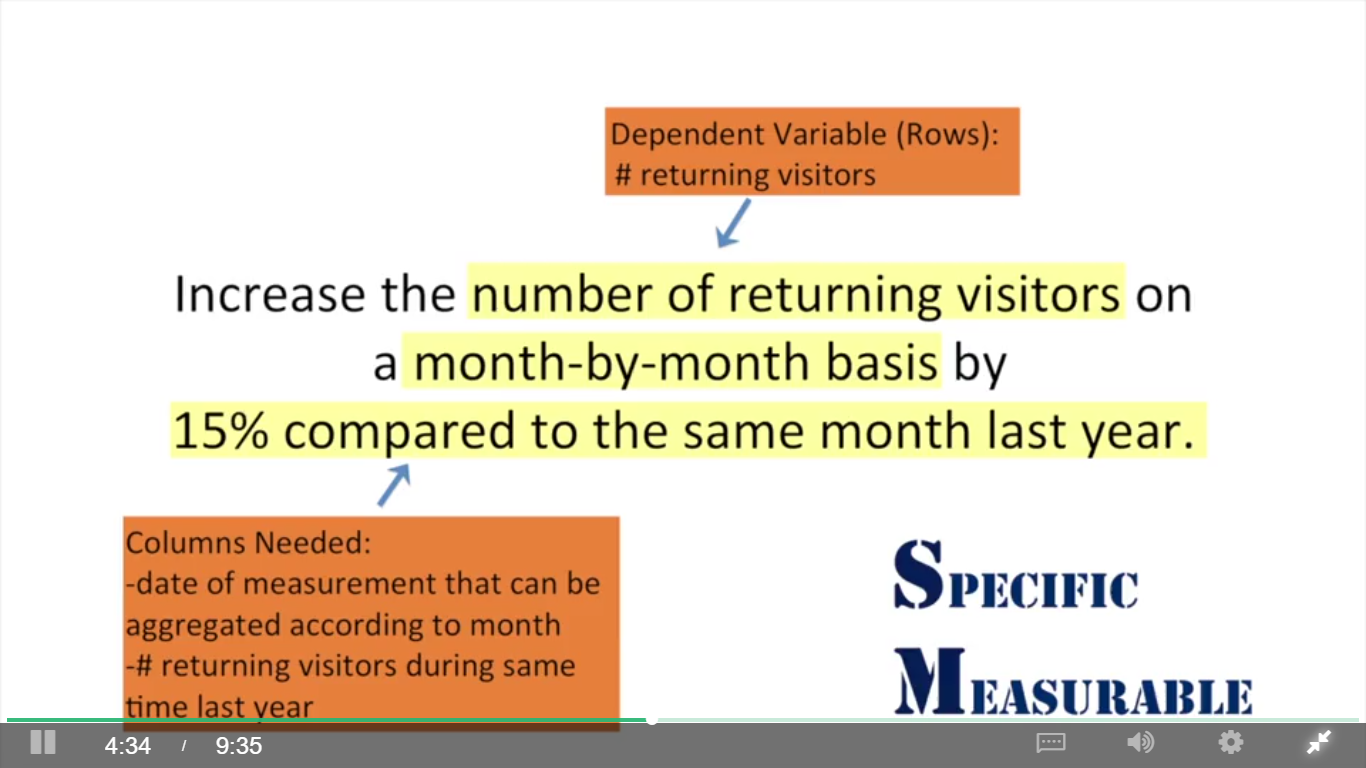
Attainable

Relevant

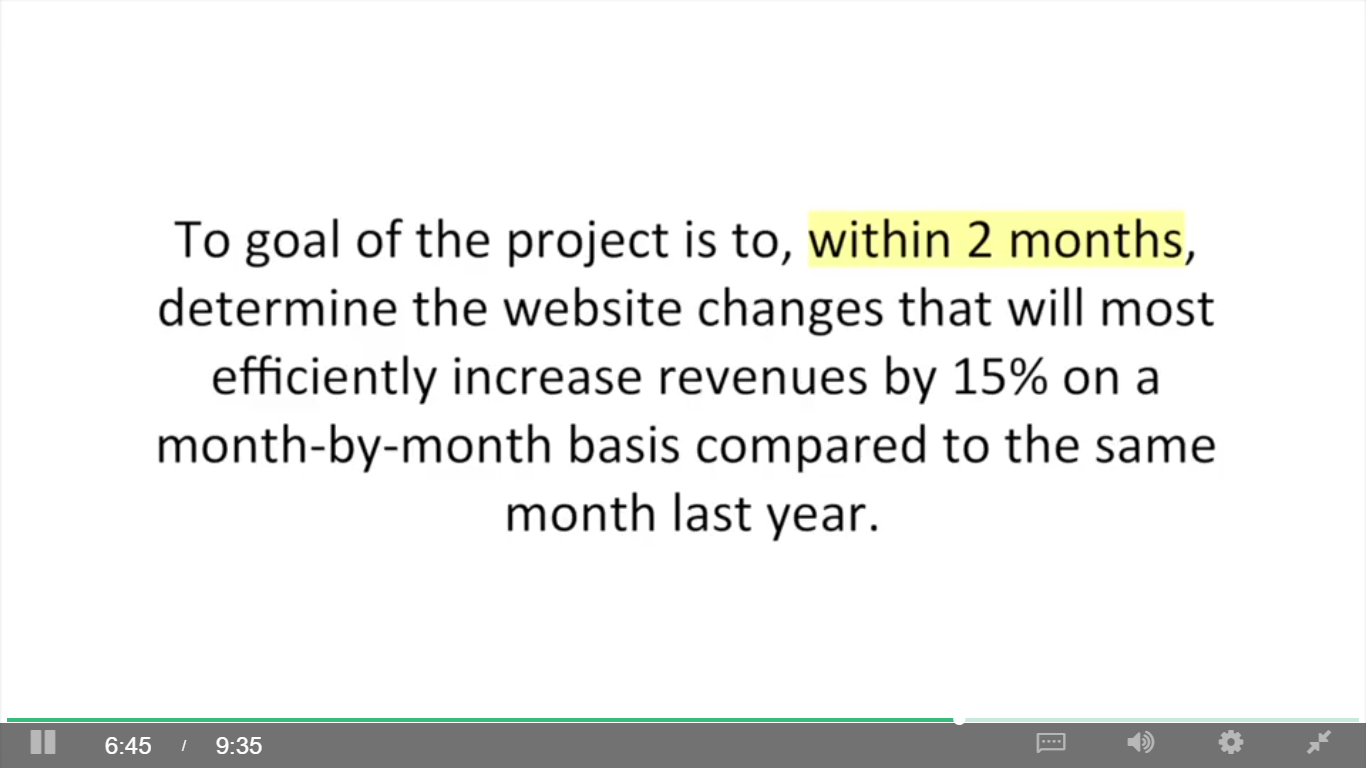
Time-bound

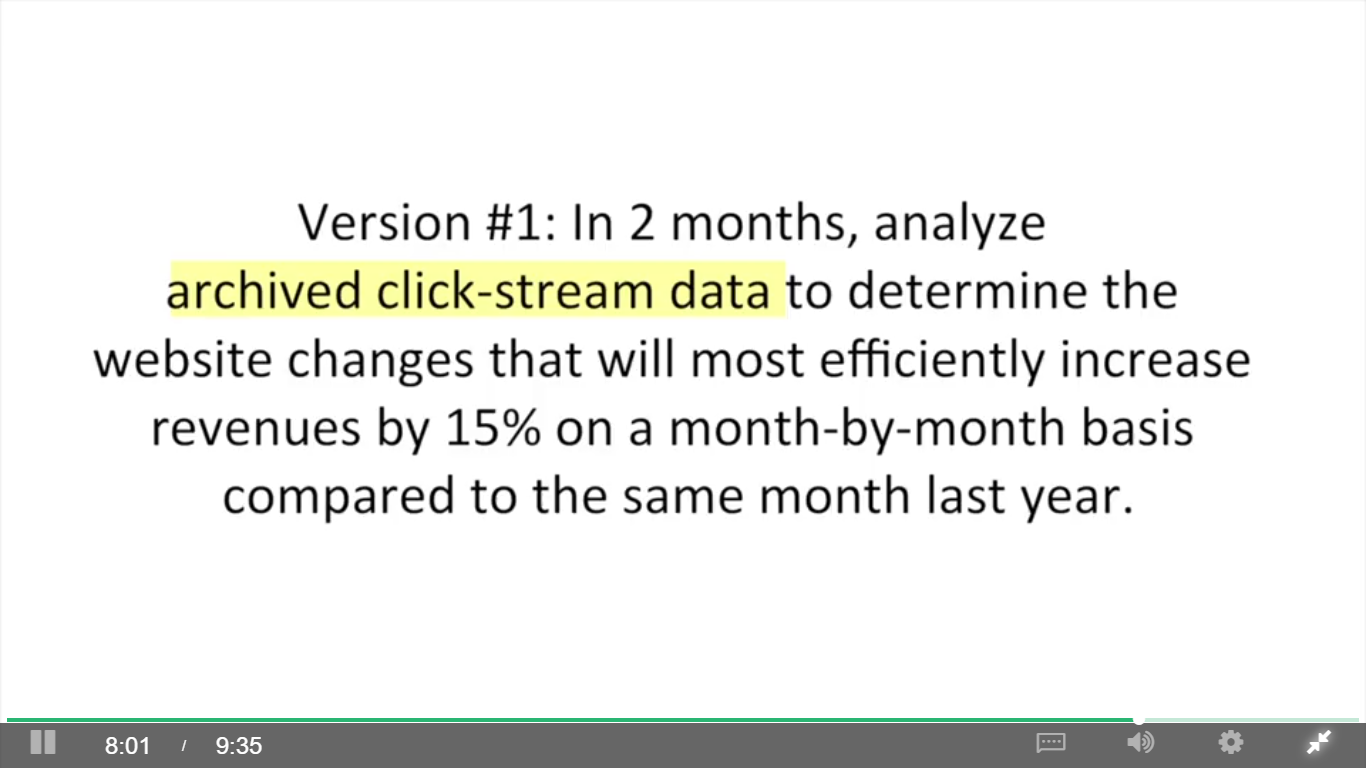
>> People aren't returning back to our site after they log in for the first time. I need you to tell us how to convert first time visitors to returning visitors.

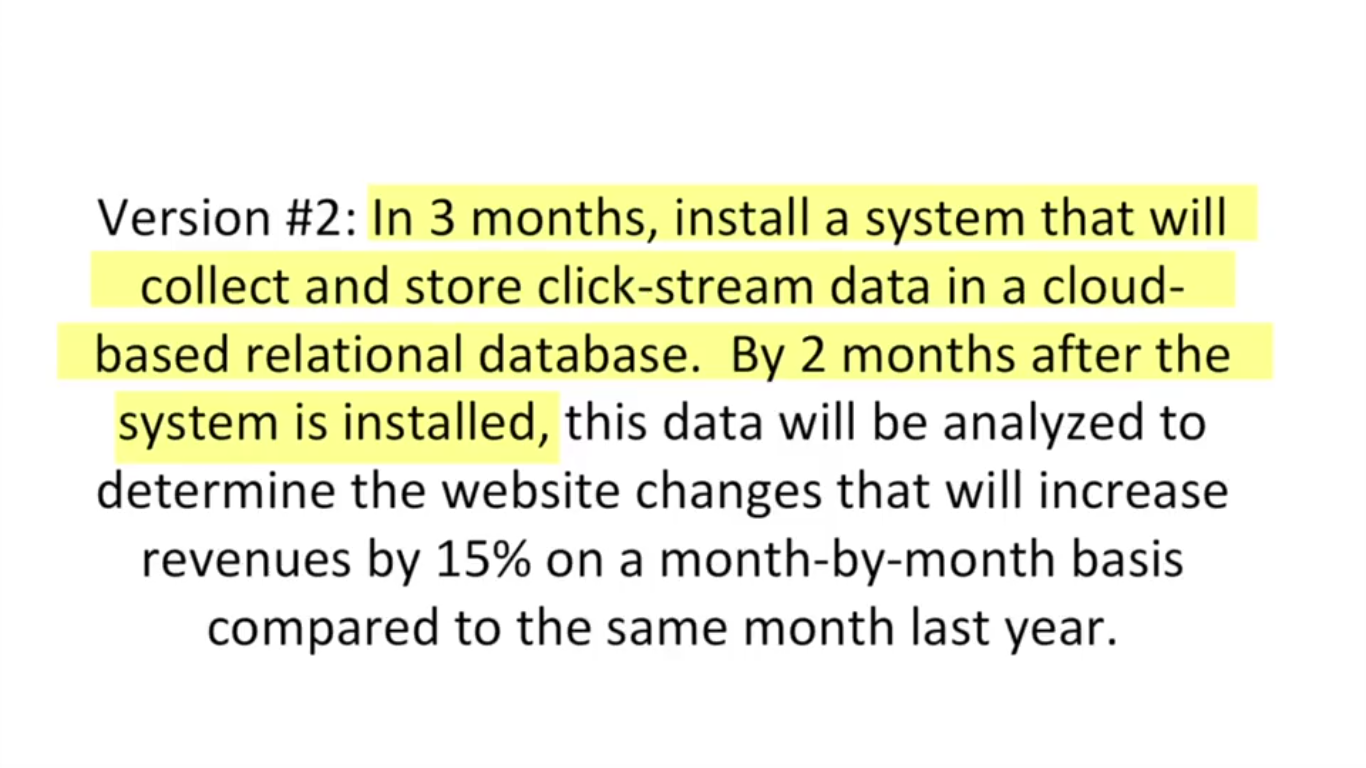




She was just making the assumption based on things she read. But the reasons profits were down was because customers weren't returning to the website. That's a good hypothesis, but in reality we have no idea whether the number of returning customers will be relevant to the drop in revenue. And if it turns out that the number of returning customers is not relevant to the drop in revenue, and your analysis is solely dedicated to the changing of the number of returning customers, your project manager and the rest of your stakeholders are not gonna be very happy.







## 1.3 Your Stakeholders

### Listening to Stakeholders During Elicitation

Methods: in person meetings, focus groups, and emails

Goal

1. Identify your key stakeholders

2. Identify independent variables to test

what has been tried before and how did it turn out? What do you think might improve this business problem?

3. determine whether stakeholders agree about problem to be solved



### Stakeholder Expectations Matter

Gartner:

Analytics continuum:

Descriptive: what happened or is happening

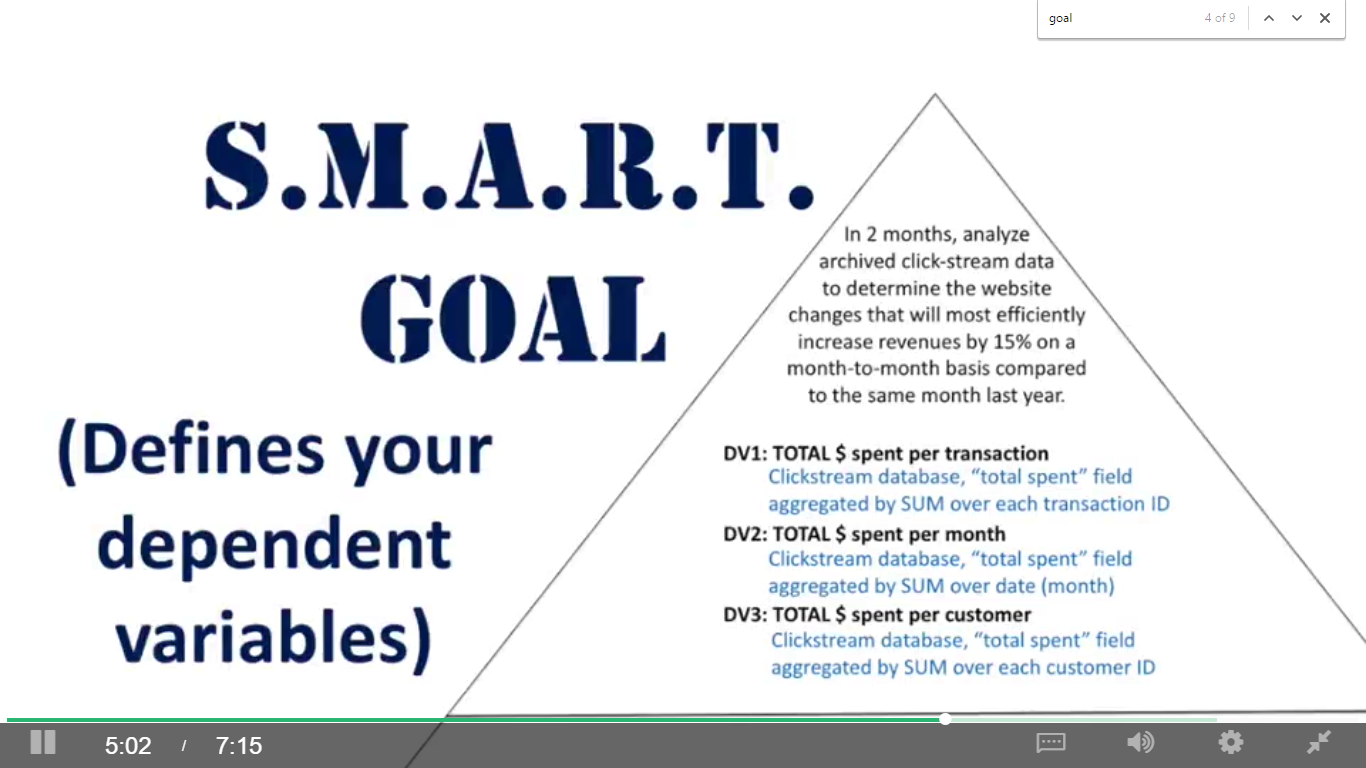
Diagnostic: why things are happening ( fraud detection)

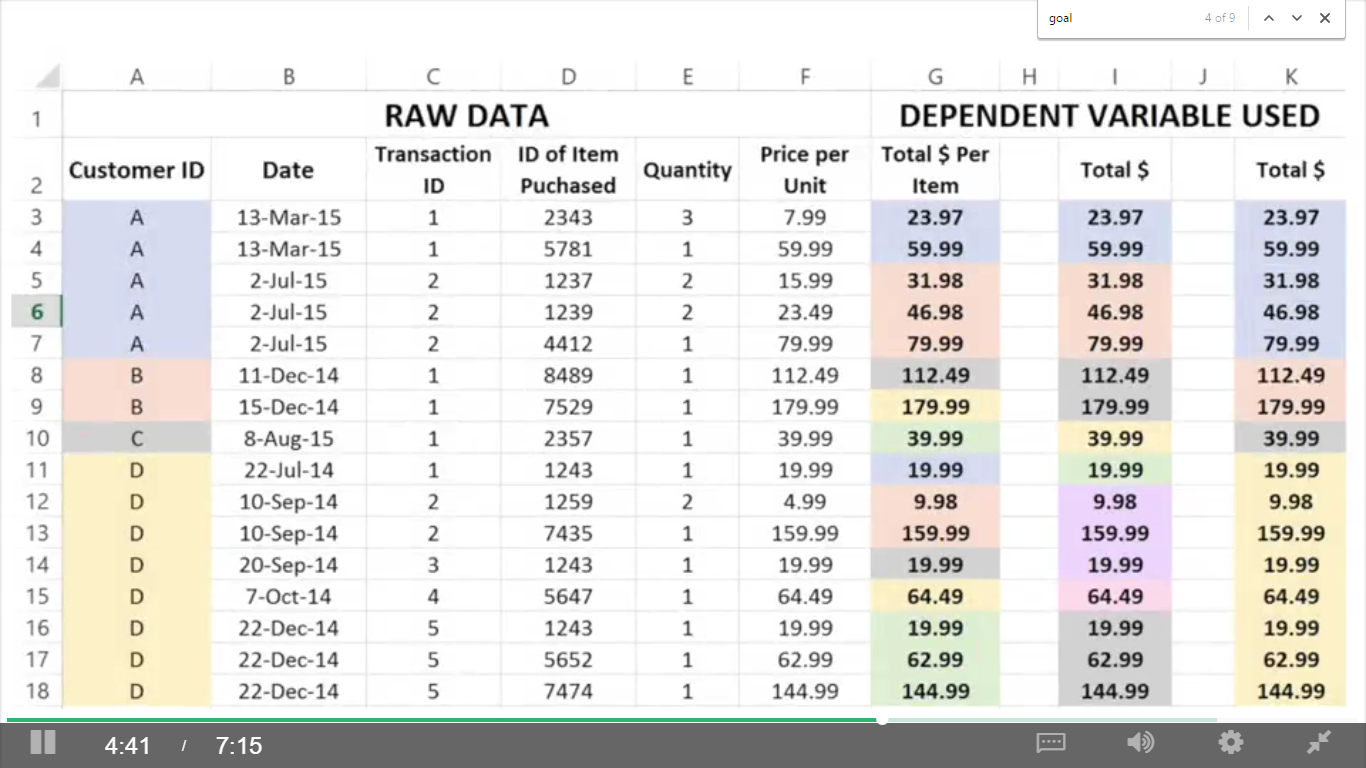
Predictive: what is going to happen (forecasting)

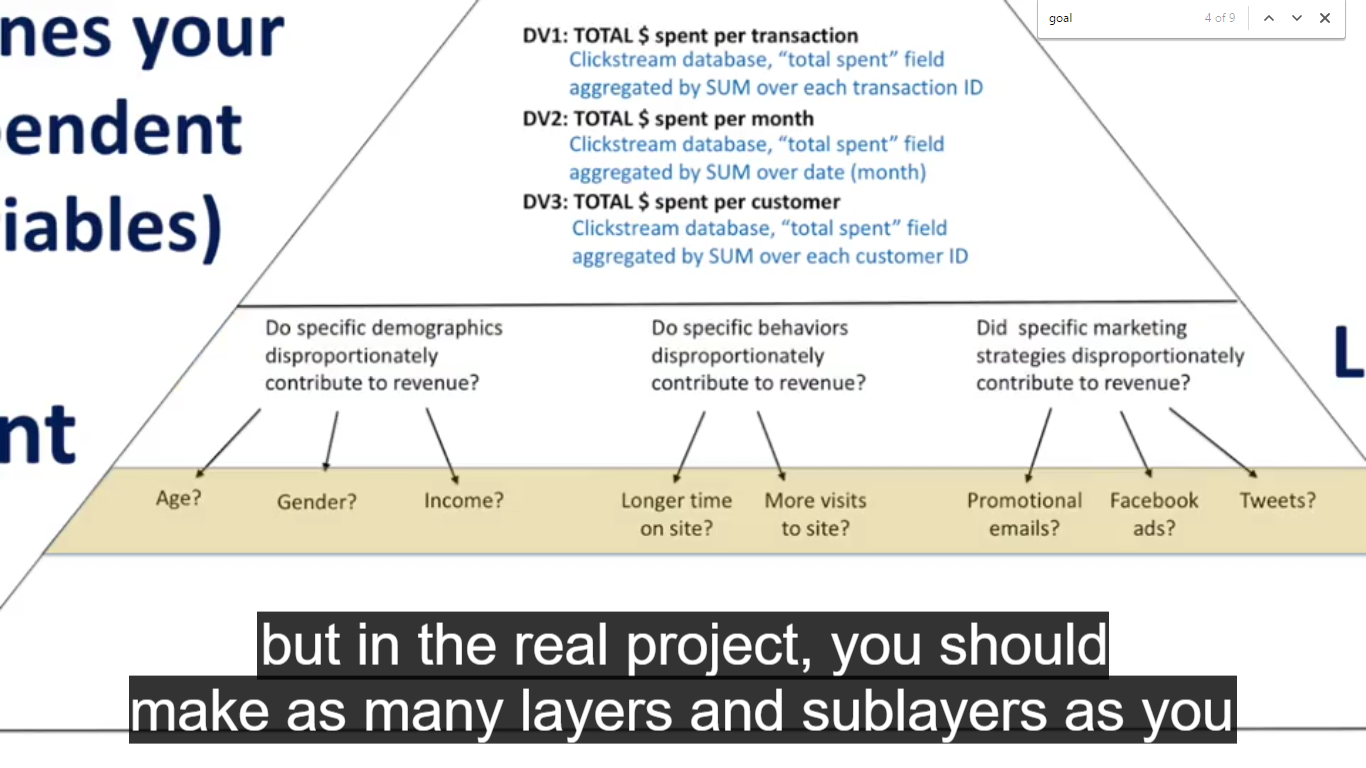
Prescriptive: giving recommendations

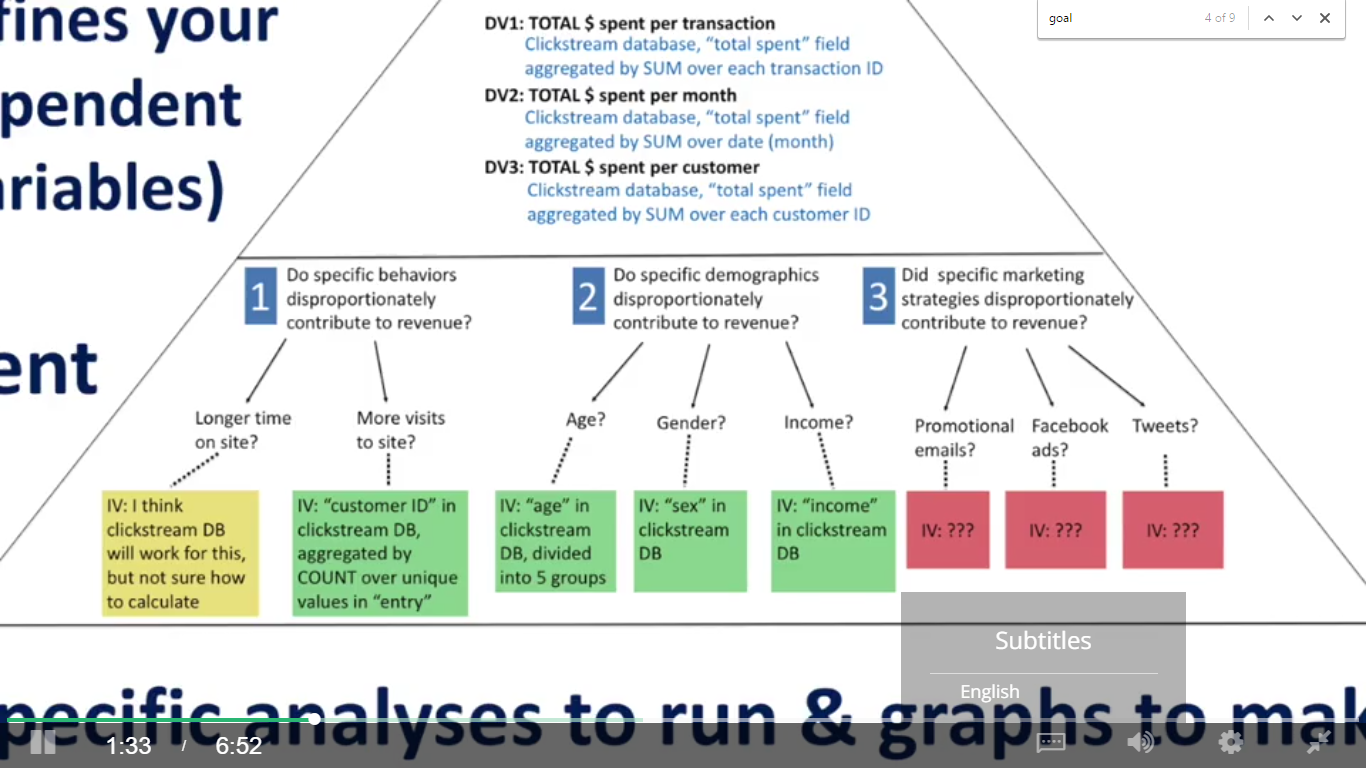
Understand the cause of stakeholders’ reluctance

## 1.4 Structure Pyramid Analysis Plans

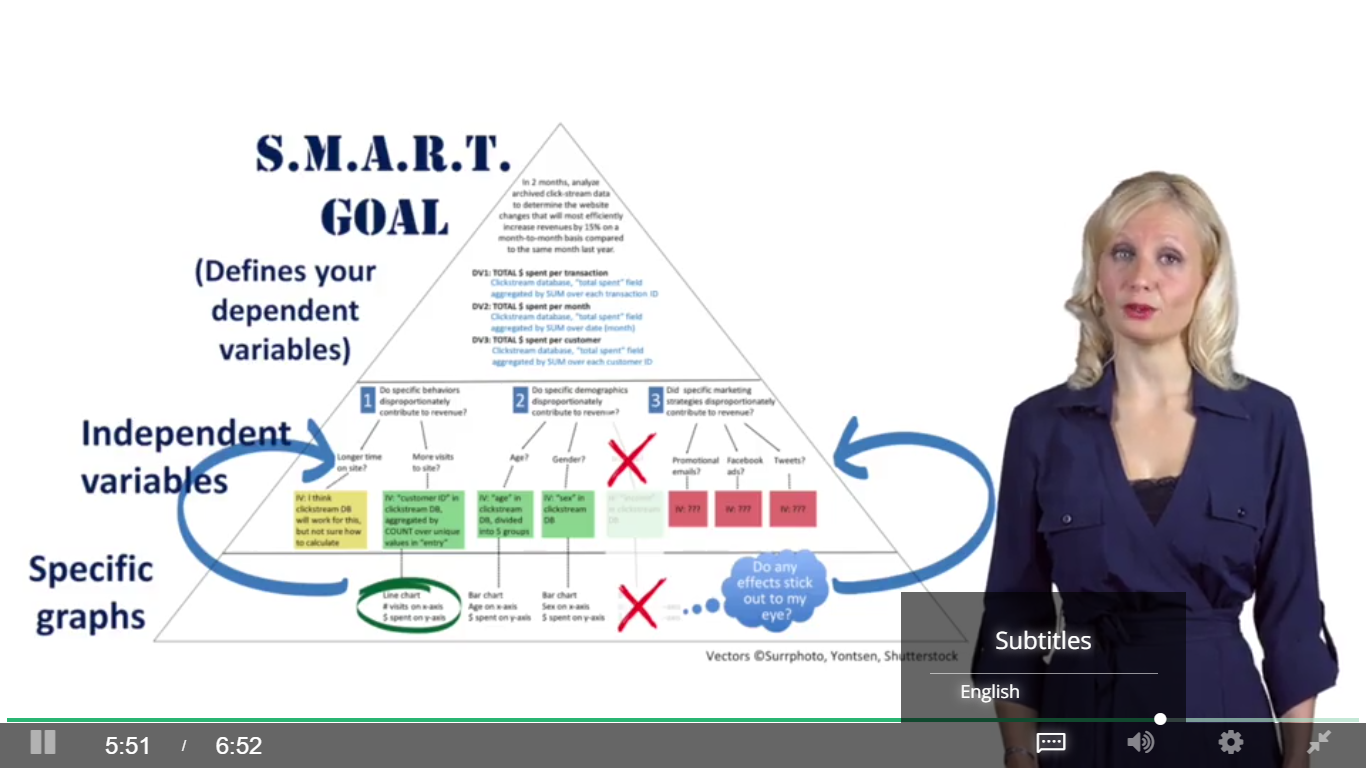








assign each category or subcategory a priority based on who suggested them, how much impact you suspect they could have, and how feasible you suspect they will be to assess.



# Week2 Data Visualization with Tableau

## 2.1 Introduction

Why Tableau?

Connect databases and import data of big sizes

Easy to draw interactive dashboards

It will help you clean your data, find patterns in your data, test your hypothesis, and tell stories with your data



## 2.2 Using Tableau to Determine How Much You Can Make as a Data Analyst

## 2.3 Let’s Get Started!

Measures: continuous variables (green)

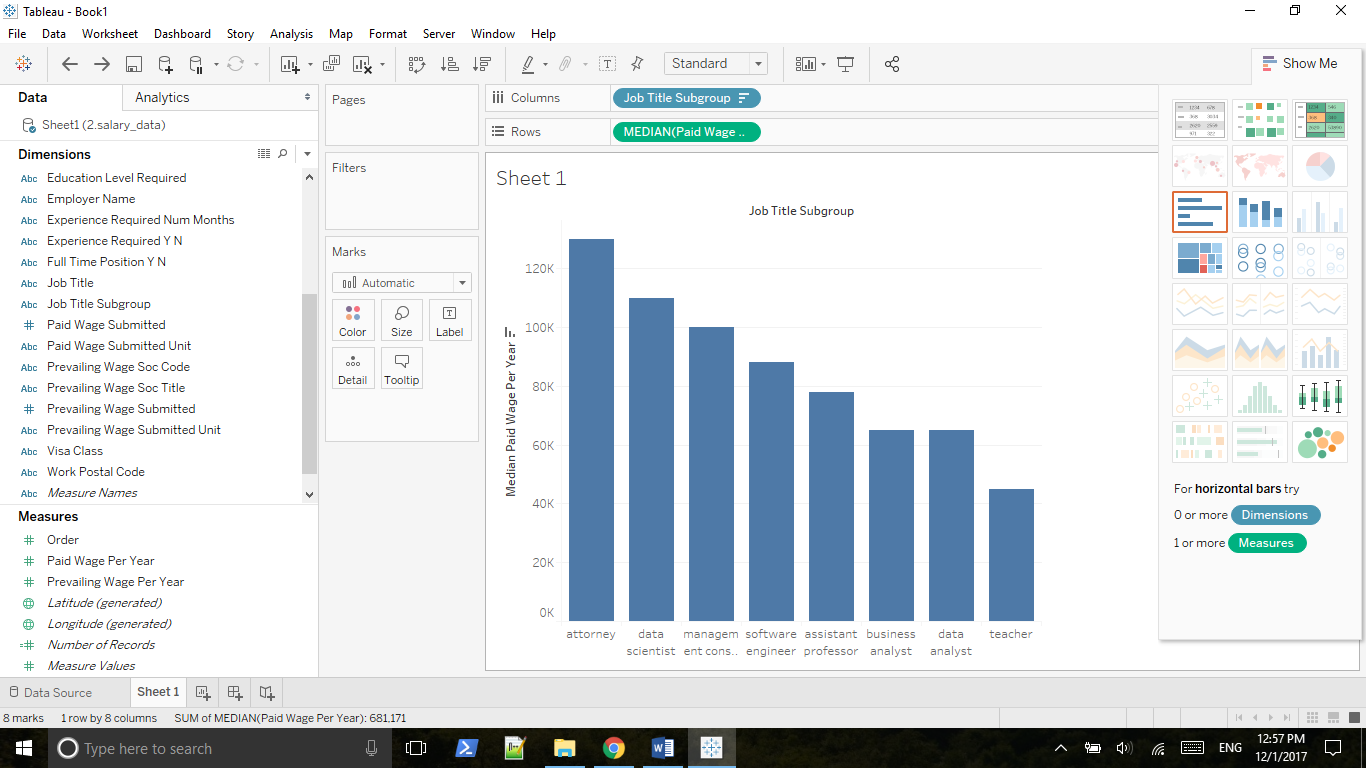
Dimensions: discreet or categorical variables (blue)

Cards and shelves

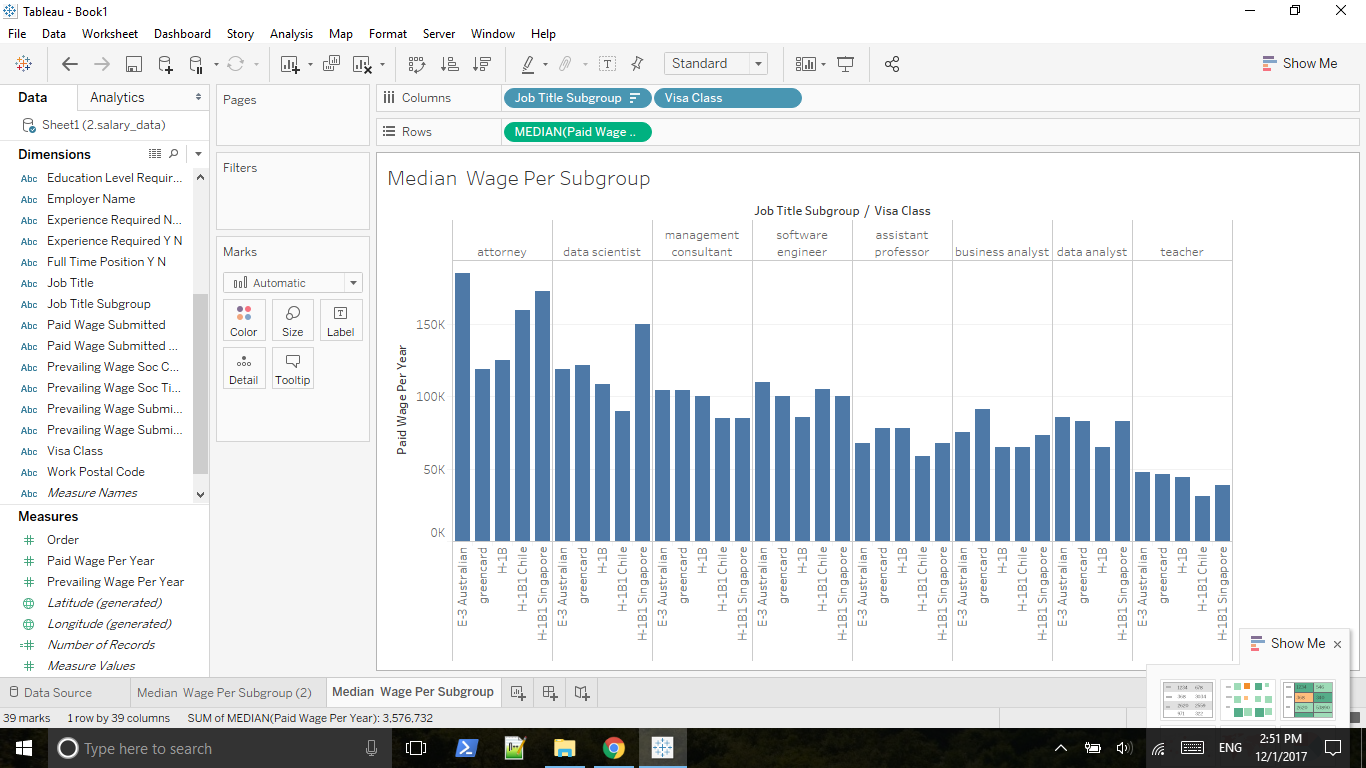
Dependent variable = paid wage per year (rows)

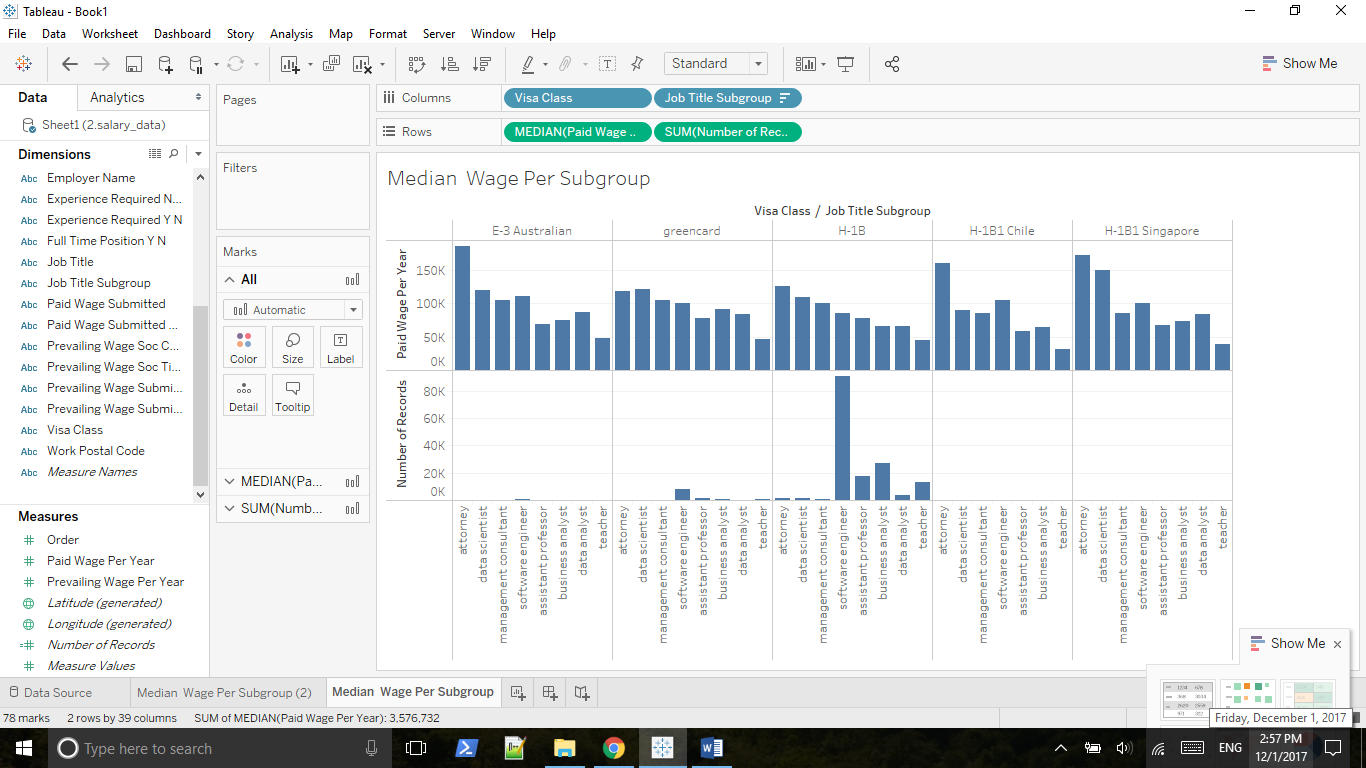
Independent variable = job title subgroup (columns)

Columns and rows, sort, switch columns and rows

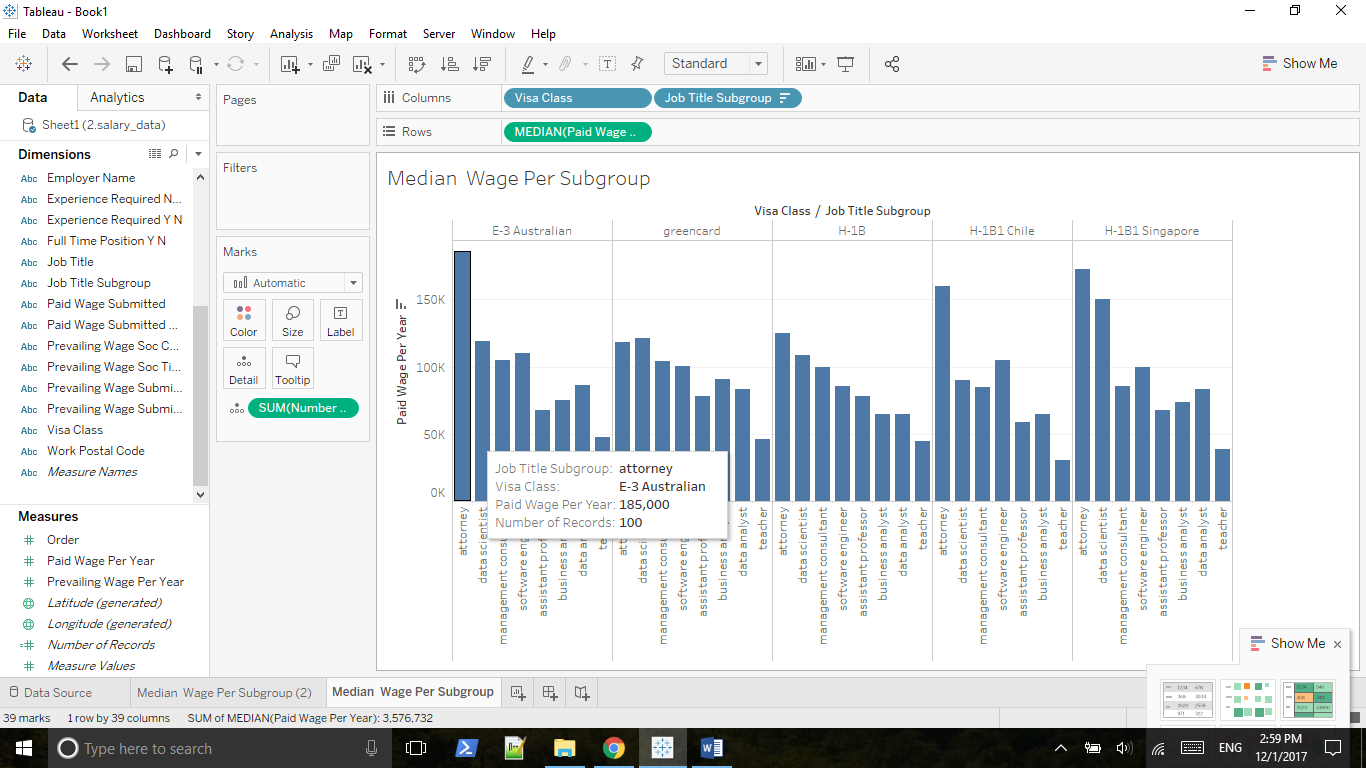


Sub dimension



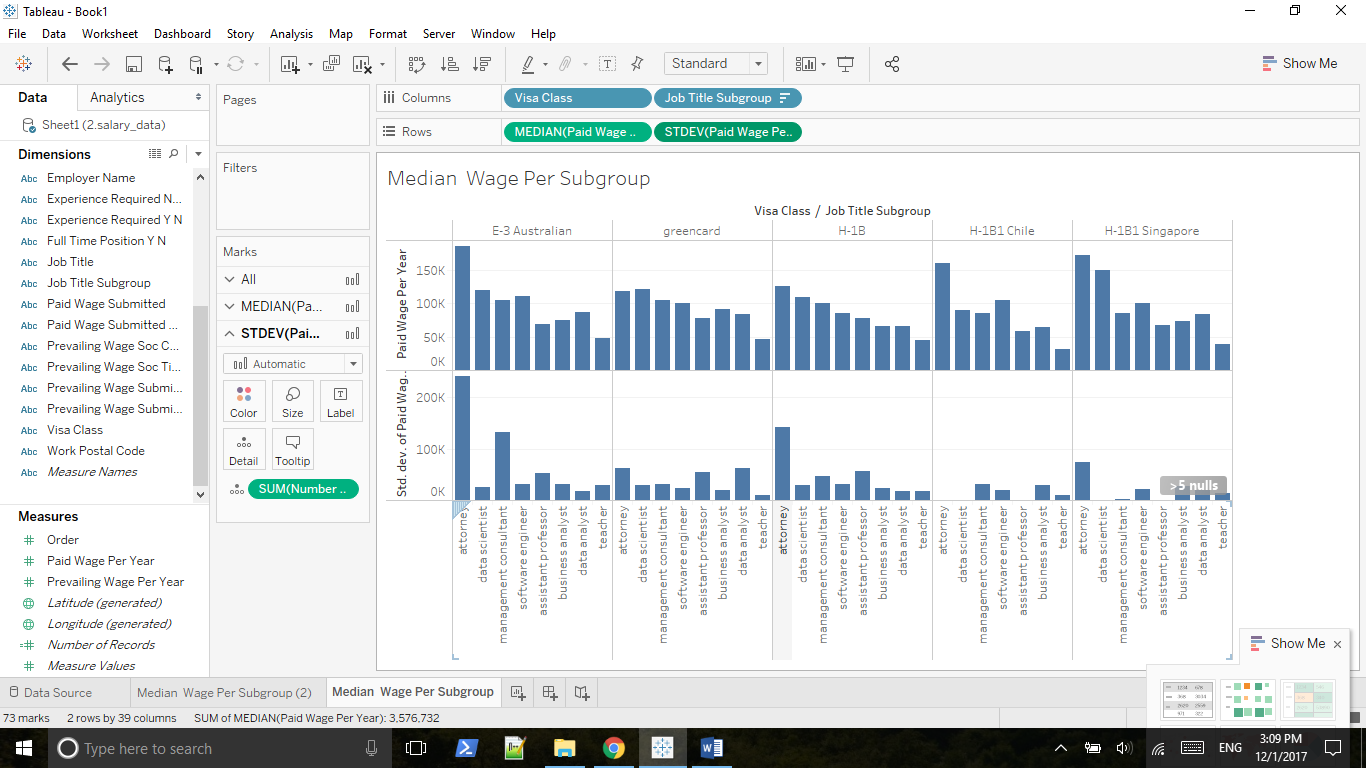


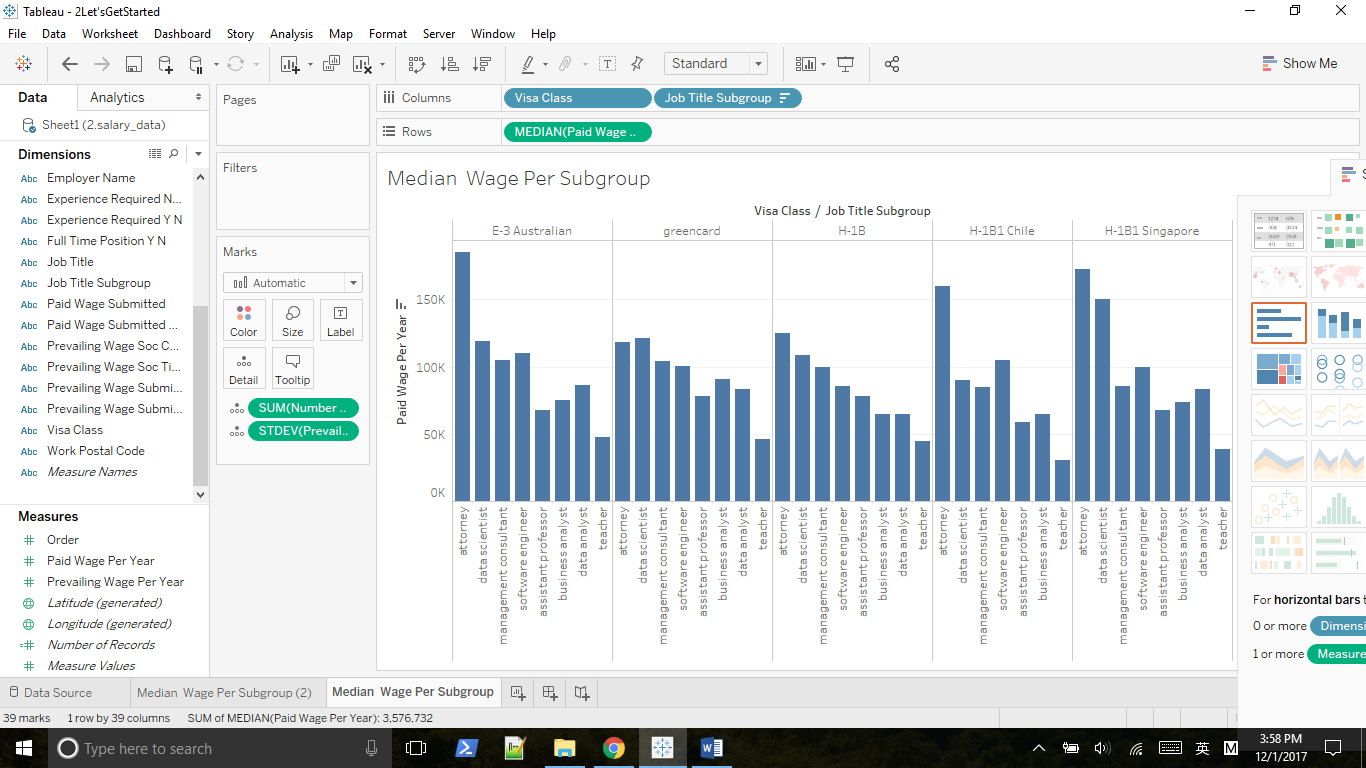
Number of records showing on the tooltip



Std. Dev (Pop) vs Std. Dev



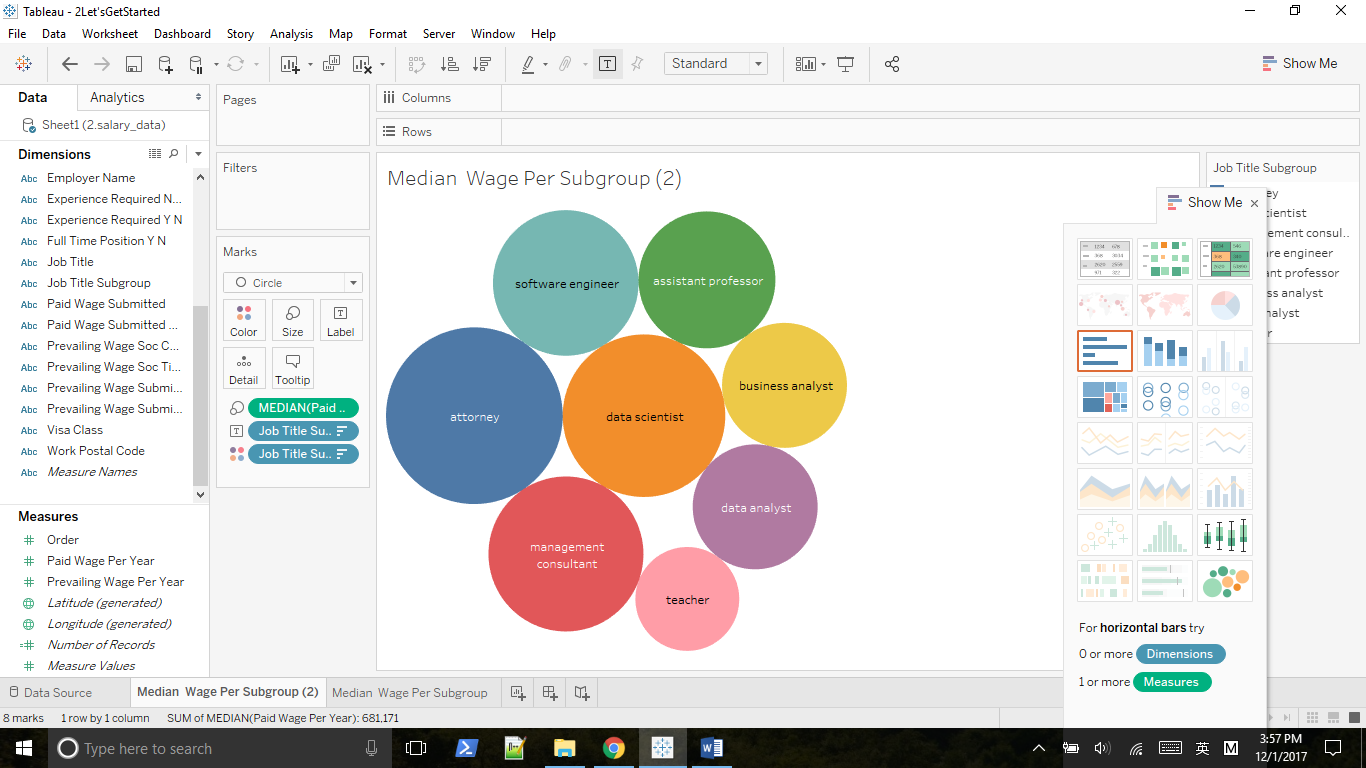




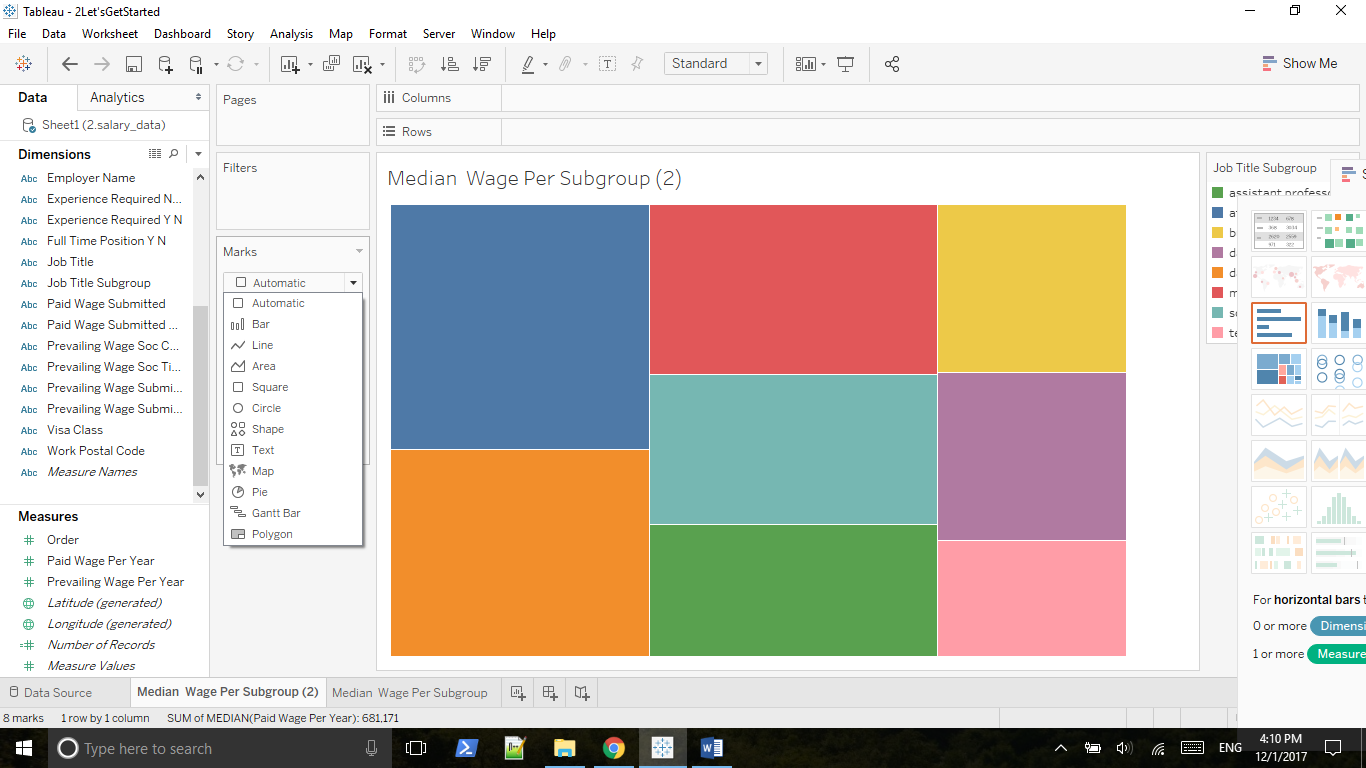
Huge std. dev implies the existence of outliers.

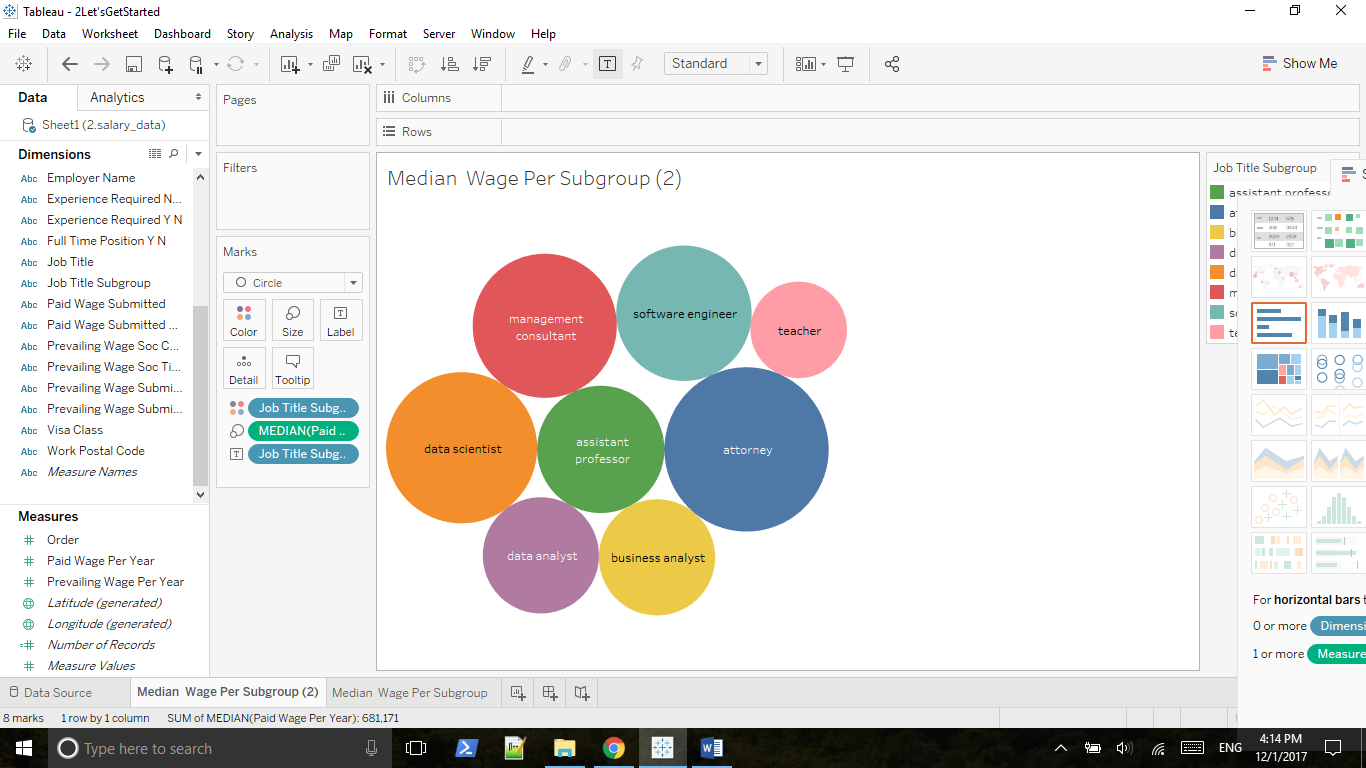
Press Shift key and move STDEV to the detail property.

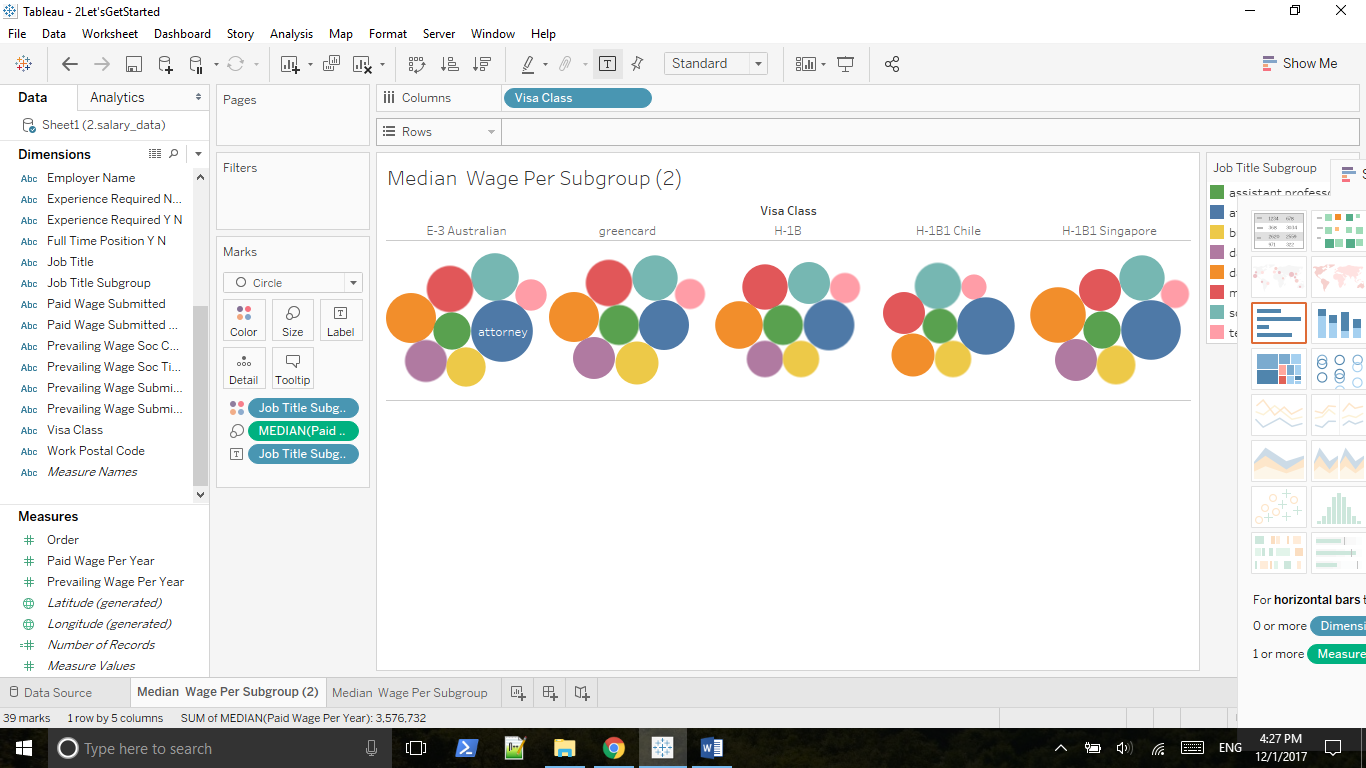
## 2.4 Working with the Marks Card



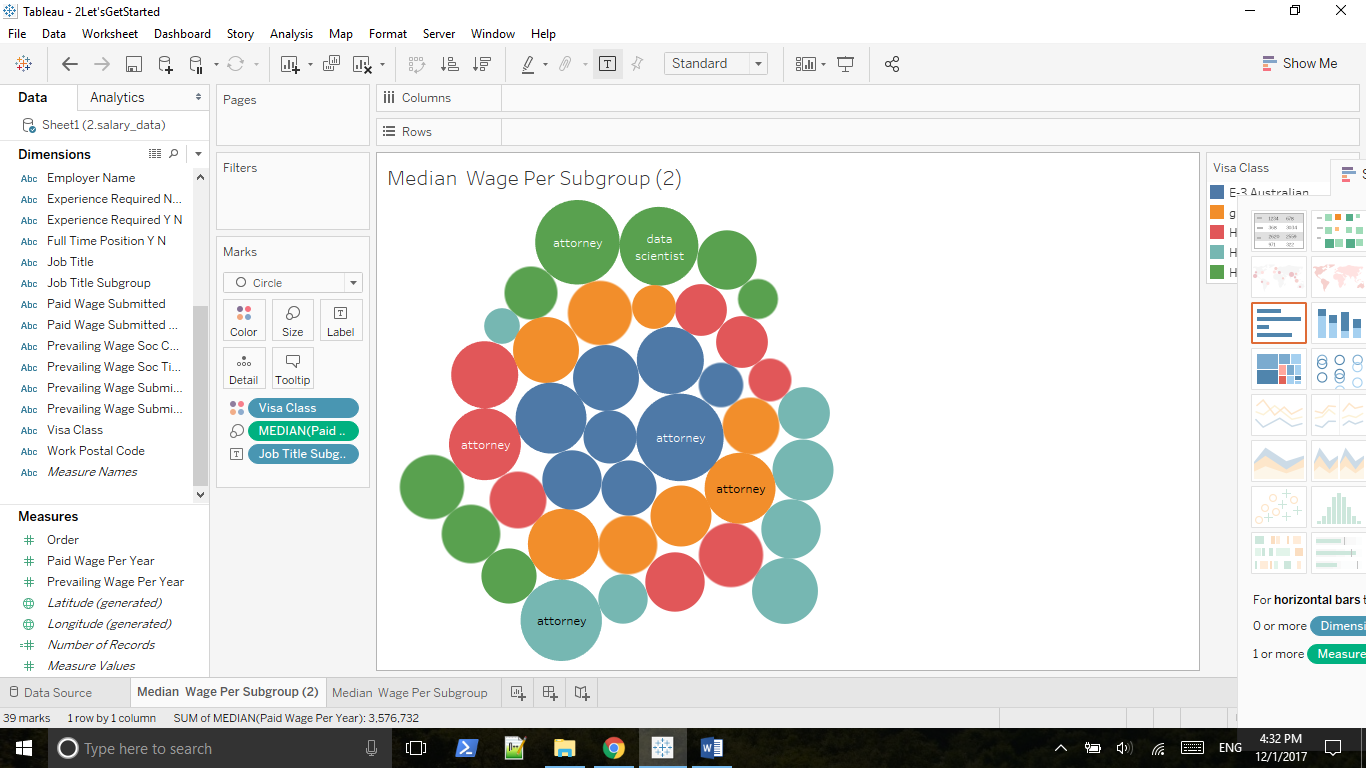
Move median to the size property



Move job title to the label property



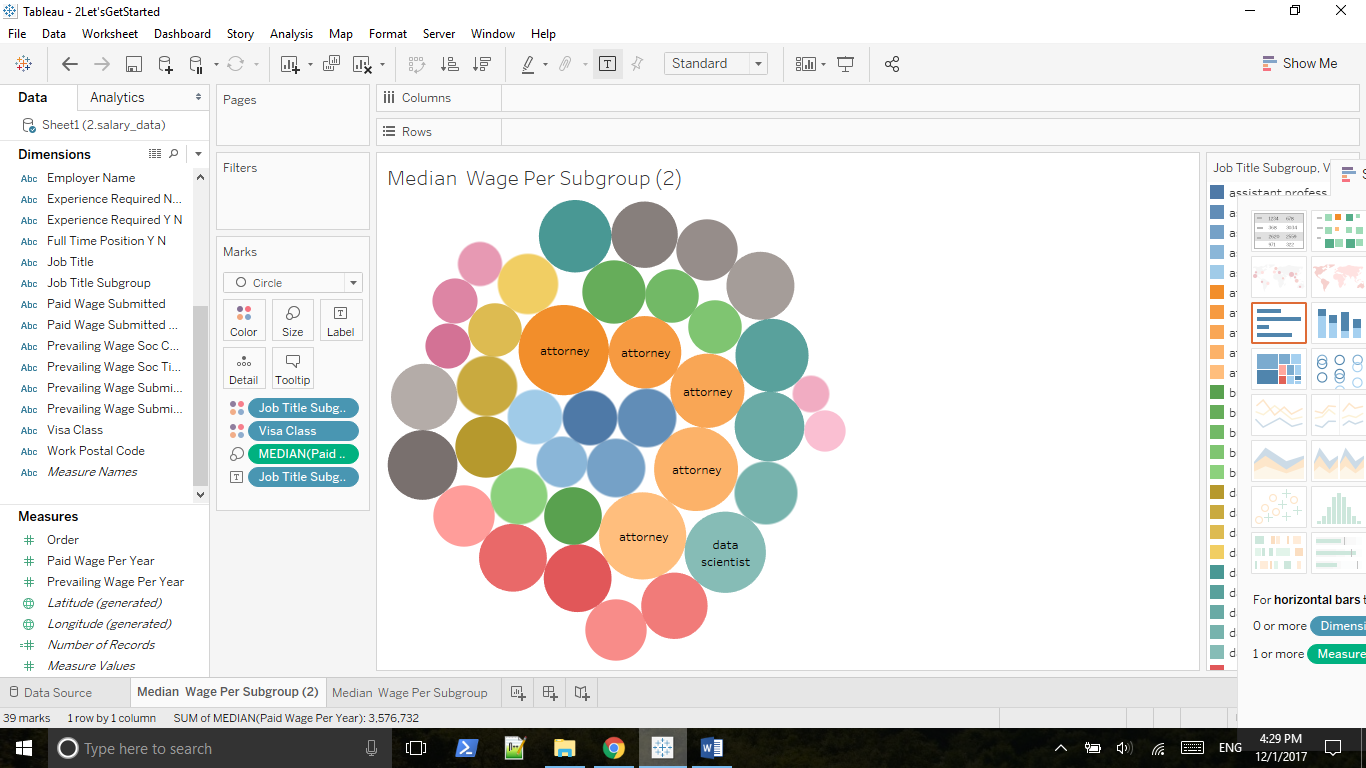
Coding visa class by color, coding the wage per year by size, coding the subgroup by label



Make label => always show

Press shift together, coding subgroup by color and coding visa class by different shade of the color

Bad interpretation, since shade usually represents a continuous variable



2.5 Outliers, Filtering and Groups