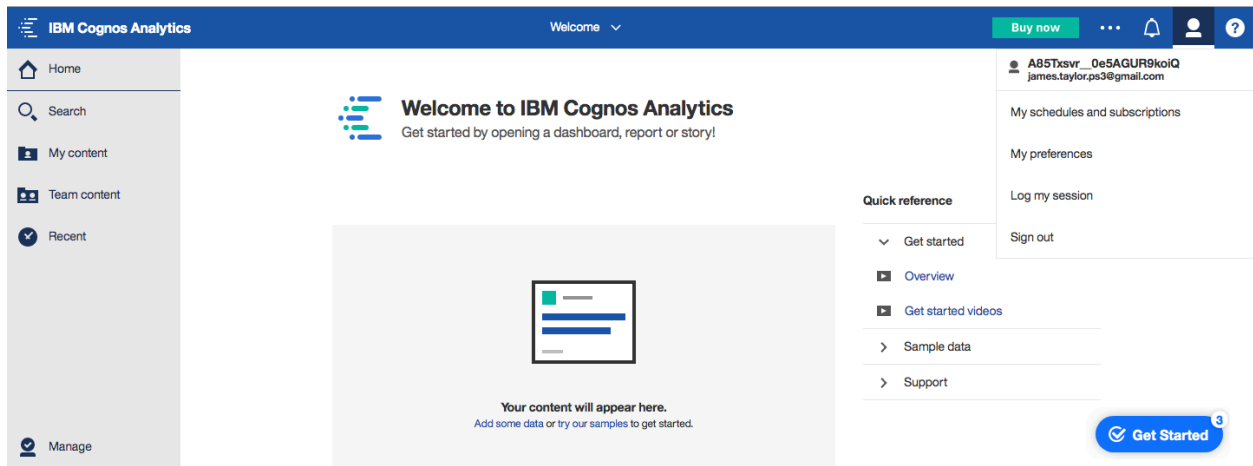


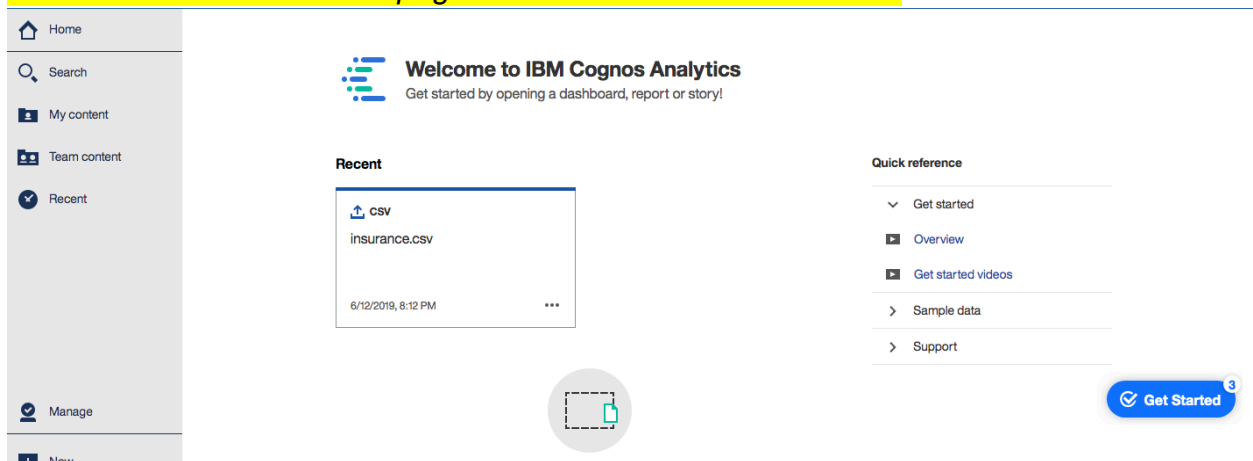
Part 1 - Account access verification

Take a screenshot of your user information. Paste the screenshot into the answer sheet



Part 2 – Upload Data

screenshot of the welcome page with the tile for insurance data



Part 3 - Data Exploration

Variable Properties

How many variables are in the insurance data?

8

What is the datatype of Age?

Integer

What is the datatype of Charges?

Decimal

What is the usage type of an Age column?

Measure

What aggregation types are available for the Smoker column?

None, Average, Count, Count Distinct, Maximum, Minimum and Total.

Explore the Relationships

What is the relationship strength between smoker and charges?

75%

What is the relationship strength between age and charges?

54%

What does the relative width of connecting line represent?

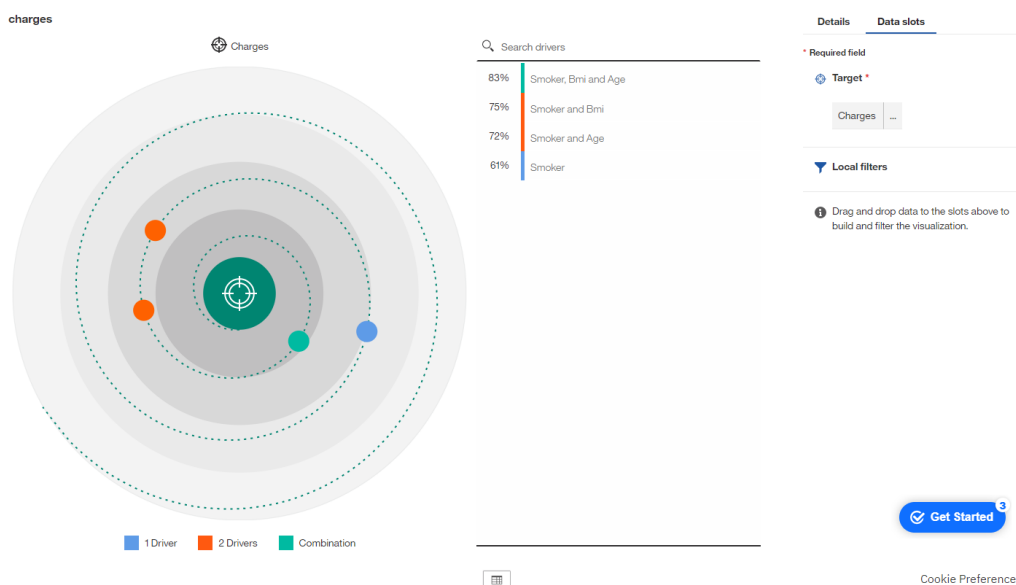
It represents the strength between the two variables.

Which variables do not have a direct relationship with charges?

Bmi and Region variables do not have a direct relationship.

Spiral Diagram

Screenshot of the spiral diagram and the list of drivers



What service satisfaction driver has the highest predictive strength? What is the predictive strength?

The combination of Smoker, Bmi and Age variables has the highest predictive strength at 83%.

What driver information do you see in the tooltip?

The combination of Smoker, Bmi and Age is a predictor of Charges

Predictive strength: 83%

Which variables in insurance data are not on the spiral diagram for charges?

Region, Children and Sex.

Summary Visualization and Using Filters

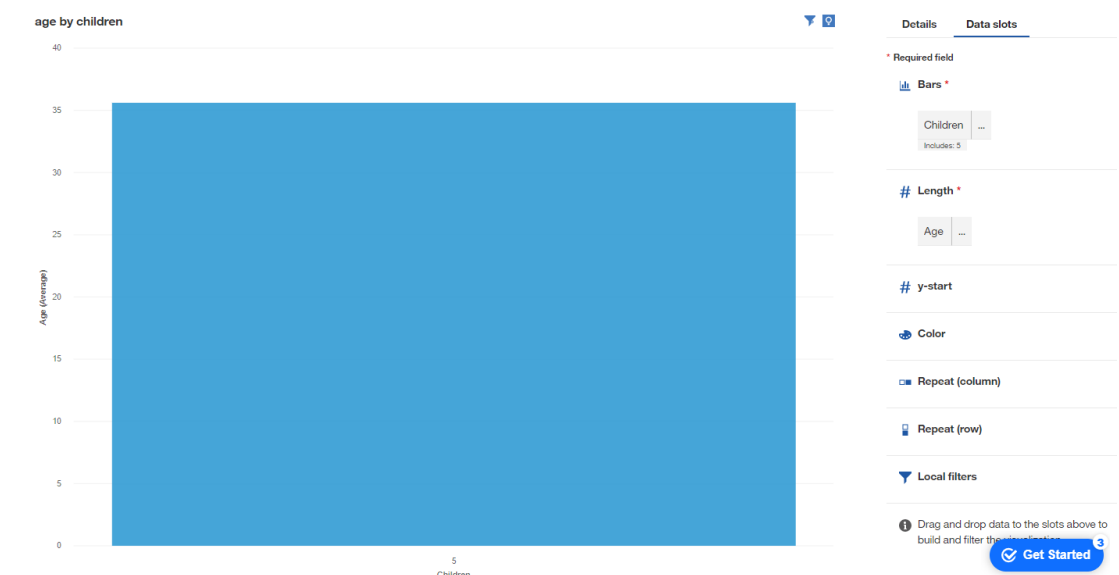
Average



Minimum



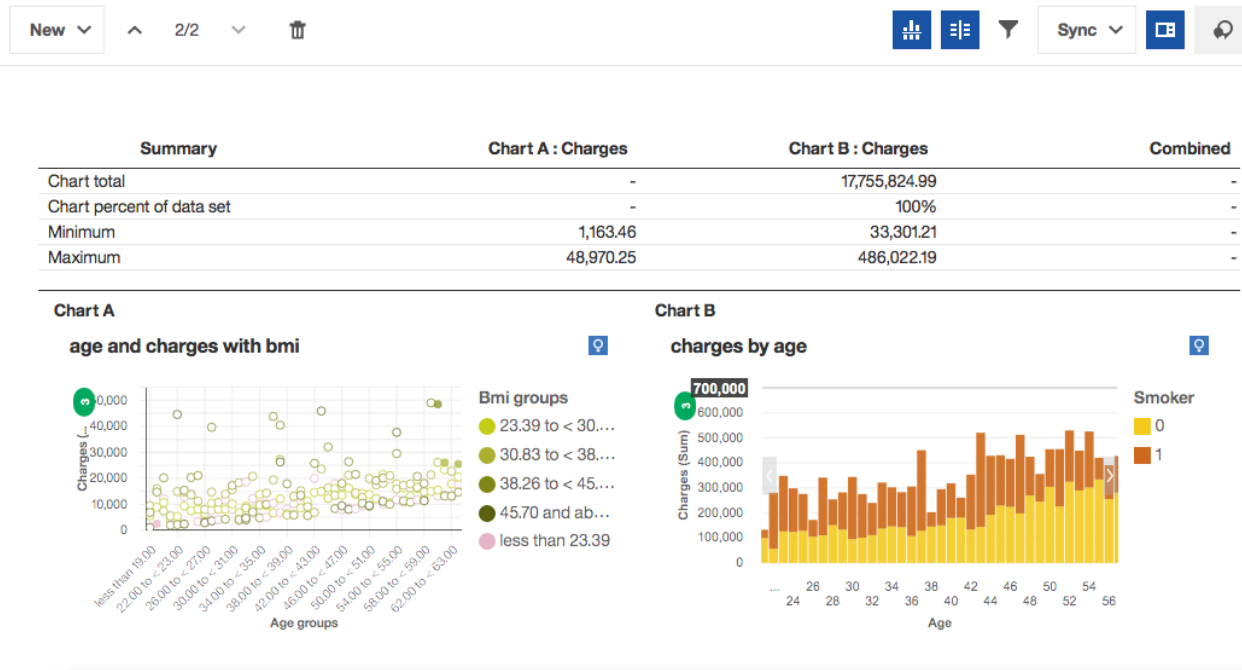
Average age of individual with 5 children.



Average age of smoker with 5 children



Comparison Visualization of Your Choice



The purpose of this comparison is to look at what may cause the increase in charges as people get older. In both of the above graphs the x-axis is age. It is clear that for every year someone gets older it is expected they will have more charges. The left graph colors the points by BMI. It is hard to see from this screenshot but generally the lowest BMI people tend to have the lowest charges irrespective of their age. This may not be surprising as weight can be correlated with overall health. The right screenshot has the proportion of smokers for each age. Like the left graphic, every additional year correlates to higher charges. The interesting part is it seems the younger groups have a higher proportion of smokers than the older groups. This could be for a couple reasons. Firstly, young people may smoke more than older people overall. If this isn't true, then either smoking itself or those who happen to smoke are live more unhealthy and riskier lives than their nonsmoking peers of the same age. It is also possible that older people used to smoke but still have health problems from their previous tobacco use. This way they would not be defined as a smoker, having quit earlier, but still have health issues.