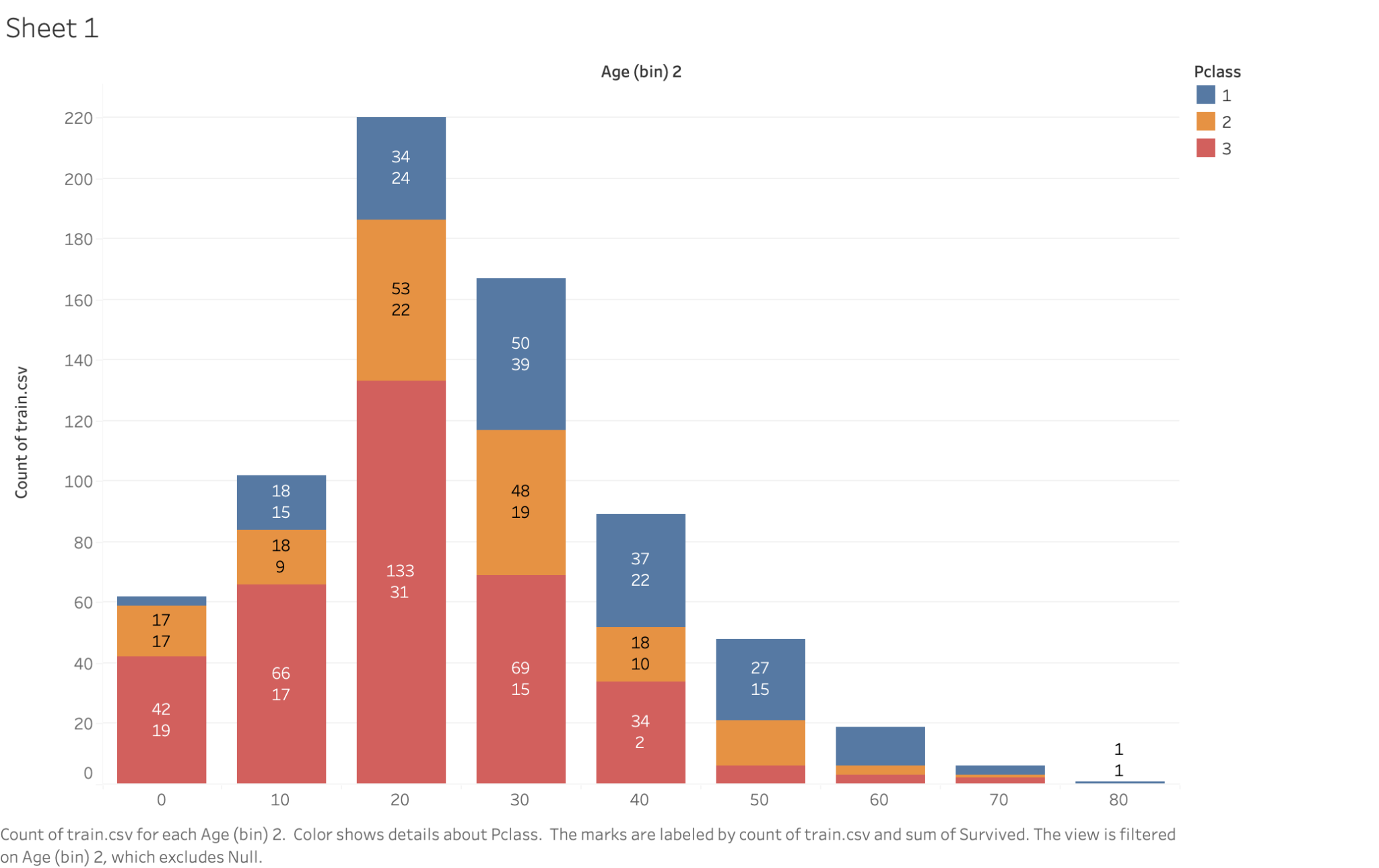
**Extra-Credit Assignment**

Visualization:



A brief description of my data set :

The dataset I used for the visualisation is the Titanic Shipwreck dataset. There are 10 variables in this dataset- survival, pclass, sex ,age, sibsp, parch, ticket, fare, cabin, embarked.

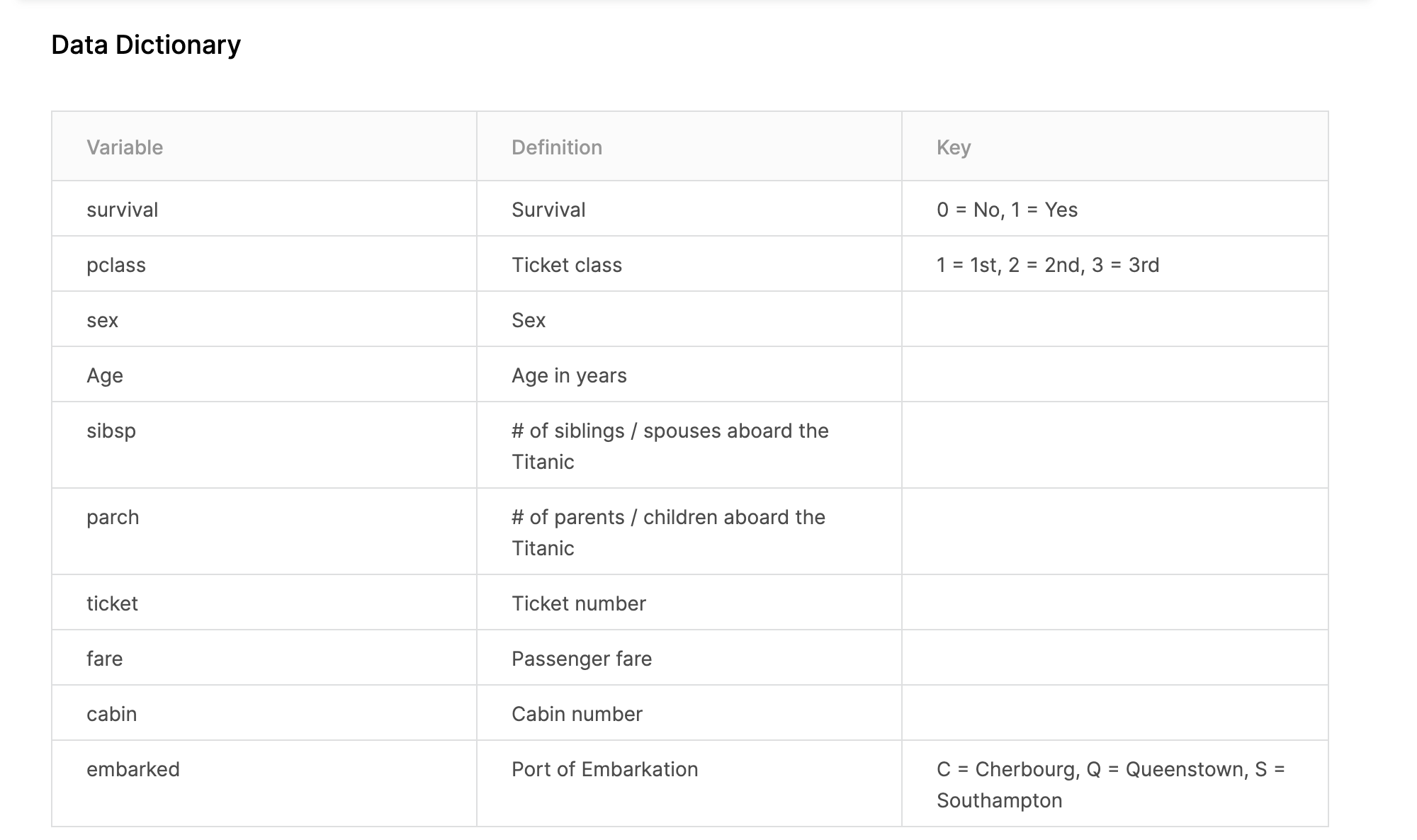
In my visualisation, I used age, pclass, survival and total number of passengers. Where Pclass is a proxy for socio-economic status (SES):

1= Upper Class

2 = Middle Class

3 = Lower Class

Below in the figure are the definitions of the variables:



A brief description of your visualization:

I created an age bin with a size of 10. This gave an age range from 0 to 80 with an interval of 10 for each.

I then took the count of the number of total passengers on the titanic and represented it on the visualisation, I also labelled it.

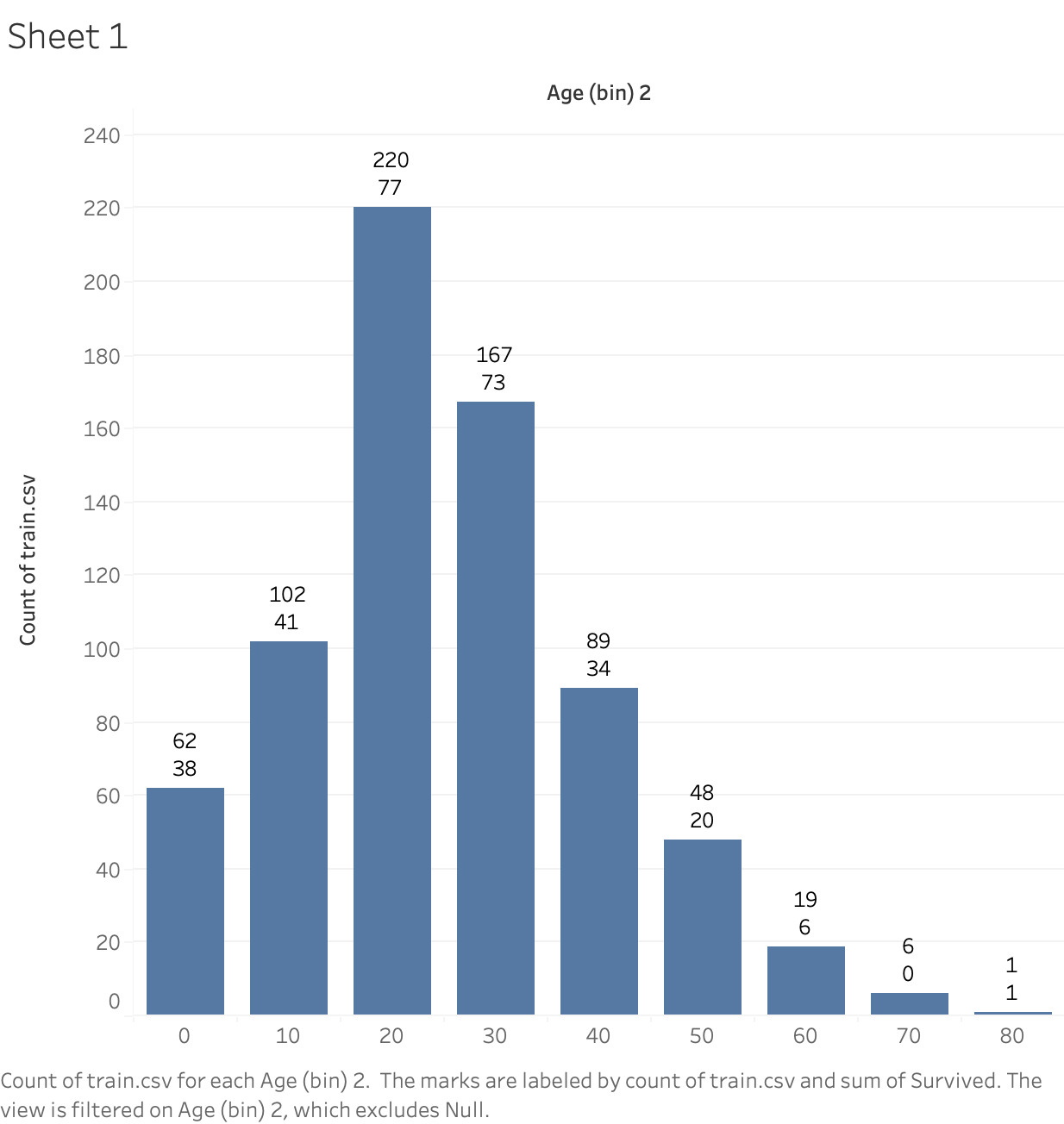
Then I took pclass and coloured them on the mark and the sum of the survived passengers were labelled which gave me my final visualisation.

The Pclass is shown by 3 colors- blue for 1, orange for 2 and pink for 3.

The first number on the bar is the count of the total number of passengers and the number below is the number of survived passengers out of them.

For example in the 0-10 bin, 42 is the total count of passengers in Pclass 3 out of which 19 survived.

Here is a visualization of only the total number of passengers in the age bins and the ones survived in them without the inclusion of pclass:



We can tell both the visualizations above that the ones with the most survival rate were passengers from age bin 0-10 and 70-80.

d) What the message of your visualization is, i.e., in words what your visualization is trying to convey, e.g., the world is getting warmer, commercial aviation is getting safer. In theory, the message should be clear from your visualization, but I want you to tell me just to be safe.

From the visualization, we can infer that the most number of passengers were from the lowest class i.e. pclass 3. We can also tell from the visualisation that the majority of the customers were in the age bracket of 20-30 (220 of them). Most of them (133) were of Pclass of 3 that is the lowest status out of which 31 survived.

I was trying to find out if a passengers Pclass had a role in determining the survival - turns out maybe it did since Pclass of 1 had the highest survival rate out of all the 3 classes. A total of 186 passengers were of pclass 1 out of which 122 survived, while only 83 survived out of 173 in pclass 2 and 85 out of 355 survived in pclass 3.