LAB1 REPORT

DATASET:

I have selected a MOVIE dataset with 15 attributes (9 categorical and 6 numerical varibles).
These were initially 5 different datasets that I downloaded from Kaggle:

1. Movie_Actors
2. Movie_Genre
3. Movie_Writer
4. Movie_Movies
5. AdditionalData

I selected some columns from each of these datasets and merged them using the imdbID attribute and randomly sampled 250 tuples for this assignment.
Following snippet shows first few rows of my final dataset.

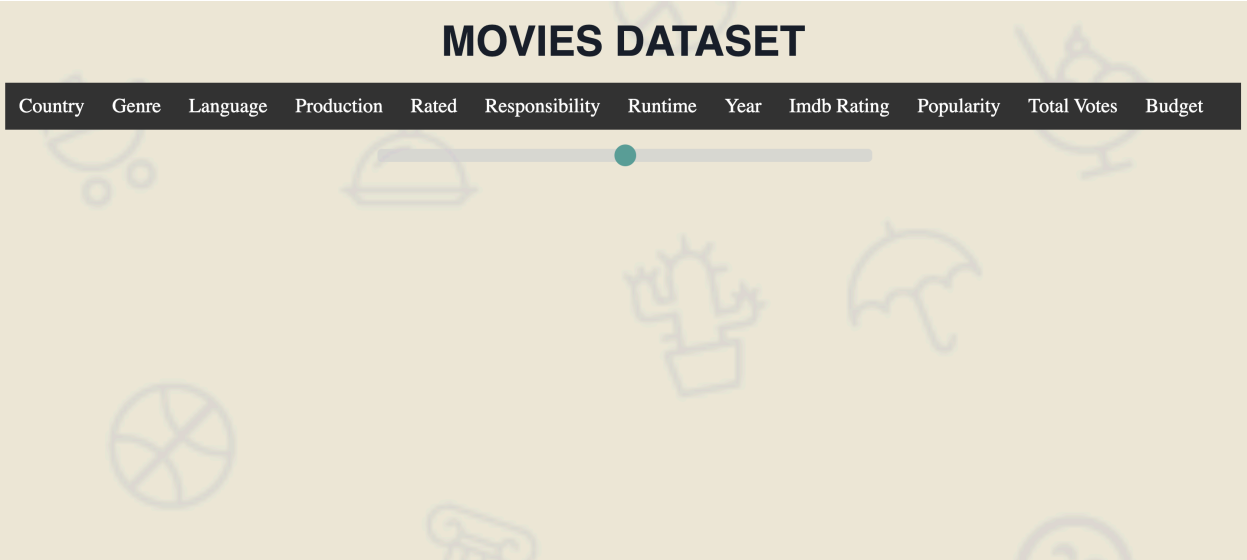
imdbID	Title	Actors	Genre	Country	Language	Production	Rated	Runtime	Year	imdbRating	Responsibility	popularity2	totalVotes	budget2
tt1174732	An Education	Cara Seymour	Drama	UK	English	Sony Classics	PG-13	100	2009	7.3	memoir	10.691	704	26096852
tt0139462	Message in a Bottle	John Savage	Drama	USA	English	Warner Bros.	PG-13	126	1999	6.2	novel	21.448	373	118880016
tt0277434	We Were Soldiers	Sam Elliott	Action	USA	English	Paramount Pictures	R	138	2002	7.2	book	14.538	846	114660784
tt2361317	Live by Night	Brendan Gleeson	Action	Spain	Spanish	Warner Bros.	R	129	2016	6.4	screenplay	8.066	751	22678555
tt0106519	Carlito's Way	John Leguizamo	Crime	USA	English	MCA Universal Home Video	R	144	1993	7.9	novels	12.644	1423	36516012
tt0091149	The Great Mouse Detective	Susanne Pollatschek	Animation	France	French	Walt Disney Pictures	G	74	1986	7.2	adaptation	33.142	744	38625550
tt2404311	The Family	John D'Leo	Comedy	USA	English	Relativity Media	R	167	2013	6.3	screenplay	12.056	1623	36894225
tt2024469	Non-Stop	Michelle Dockery	Action	UK	English	Universal Pictures	PG-13	106	2014	7	screenplay	12.239	3290	222809600
tt0094678	Arthur 2: On the Rocks	Geraldine Fitzgerald	Comedy	USA	English	Warner Bros.	PG	113	1988	4.3	characters	5.252	64	14681192
tt0086993	The Bounty	Edward Fox	Action	UK	English	MGM	PG	132	1984	7.1	book	8.415	205	8613462
tt1961175	American Assassin	Chris Webster	Action	USA	English	Lionsgate	R	177	2017	4.6	screenplay	13.803	1572	66167951
tt0084917	The World According to Garp	John Lithgow	Comedy	USA	English	Warner Bros.	R	136	1982	7.2	novel	7.816	141	29712172
tt0116136	DragonHeart	Dina Meyer	Action	USA	English	MCA Universal Home Video	PG-13	103	1996	6.4	story	44	1019	115267375

IMPLEMENTATION:

Server: I have used XAMPP server for this assignment.
I have used **HTML**,**CSS** and **Javascript** in this assignment.

1. In the html file(**lab1.html**), I have added all the components of the web page i.e the heading, navigation bar and the slider.
2. In the CSS file(**index.css**), I have done the formatting and styling of the components added on the web page.
3. In **movie.js** I have made different functions for each attribute and using these functions I send the required attribute to the bar or hist function.
4. In **hist.js** implemention of the histogram is added.
5. Similarly **graph.js** file contains the bar chart implementation .

The following snippet is my basic website design:

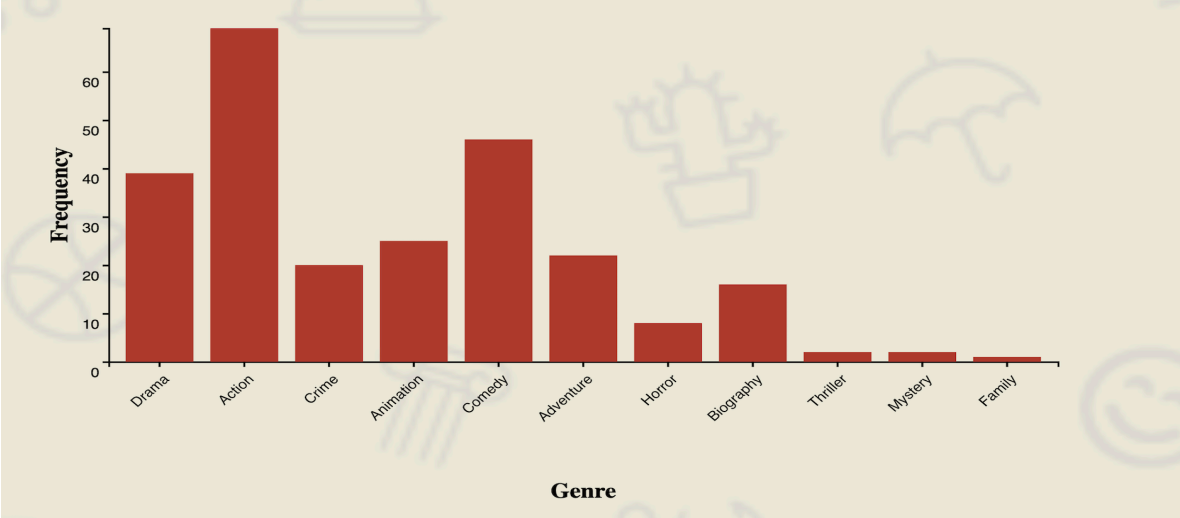


I have used a navigation bar to select any attribute of your choice. On clicking any option from the bar , it will display the chart of that particular attribute.

CATEGORICAL VARIABLES:

For each of 6 the categorical attributes(Country,Genre,Language,Production,Rated,Responsibility) ,I have created a bar chart with the attribute on the x axis and its frequency on the y-axis.

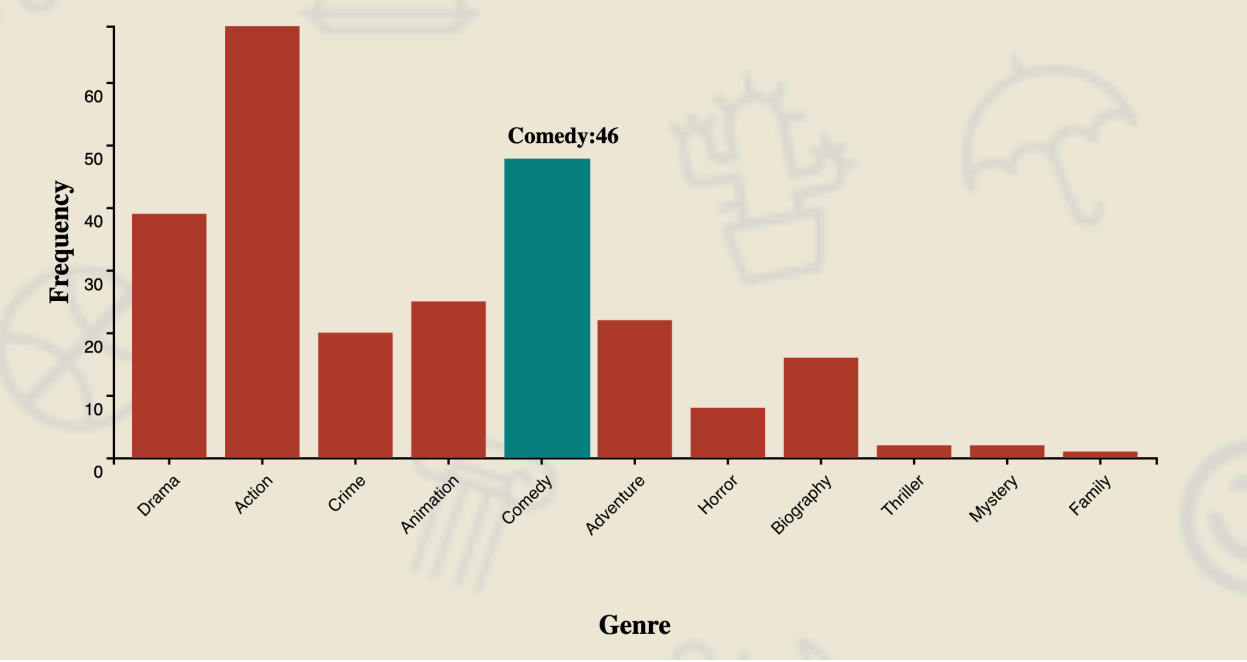
For example,
Genre is divided into 11 categories hence 11 bars with its count on the y axis.



Functionalities in the bar chart are :

1. On mouse-over there are 3 functionalities:

   a. The value of the bar graph is displayed on the top of the cursor
   b. The bar becomes wider and higher
   c. The color of the bar changes.



For implementing this part I have used tooltip:
On mousemove ,it takes the current position of x and y as xPos and yPos , and add the text (Value) to the tooltip and also increases the width and height of the bar and changes the color.
On mouseout it is reset to the original size and color.

The following figure is the code snippet of this implementation:

```
98          .on("mouseover",function(){
99              tooltip.style("display",null)
100         })
101         .on("mouseout", function(d){
102             tooltip.style("display","none")
103             d3.select(this)
104     //          // .transition().duration(500)
105             .attr("x",x(d.x0) )
106             .attr("y", y(d.length) )
107             .attr("width",x(d.x1) -x(d.x0) )
108             .attr("height",height-y(d.length) )
109             .style("fill","#B03A2E");
110         })
111         .on("mousemove",function(d){
112             var xPos=d3.mouse(this)[0]-30;
113             var yPos=d3.mouse(this)[1]-55;
114
115             tooltip.attr("transform","translate(" + xPos + ","+ yPos +")")
116             tooltip.select("text").text(d.length);
117
118              d3.select(this).attr('class','bar')
119     //          // .transition().duration(500)
120             .attr("x",x(d.x0) -5)
121             .attr("y", y(d.length)-15)
122             .attr("width",x(d.x1) -x(d.x0) +10 )
123             .attr("height",height-y(d.length)+15 )
124             .style("fill","#008080");
125
126
127         });
128
129         var tooltip = svg.append("g")
130             .attr("class",tooltip)
131             .style("display","none")
132             .text("");
133
134         tooltip.append("text")
135             .attr("x",15)
136             .attr("dy","1.2em")
137             .style("font-size","1.5em")
138             .style("font-weight","bold");
```

NUMERICAL ATTRIBUTES:

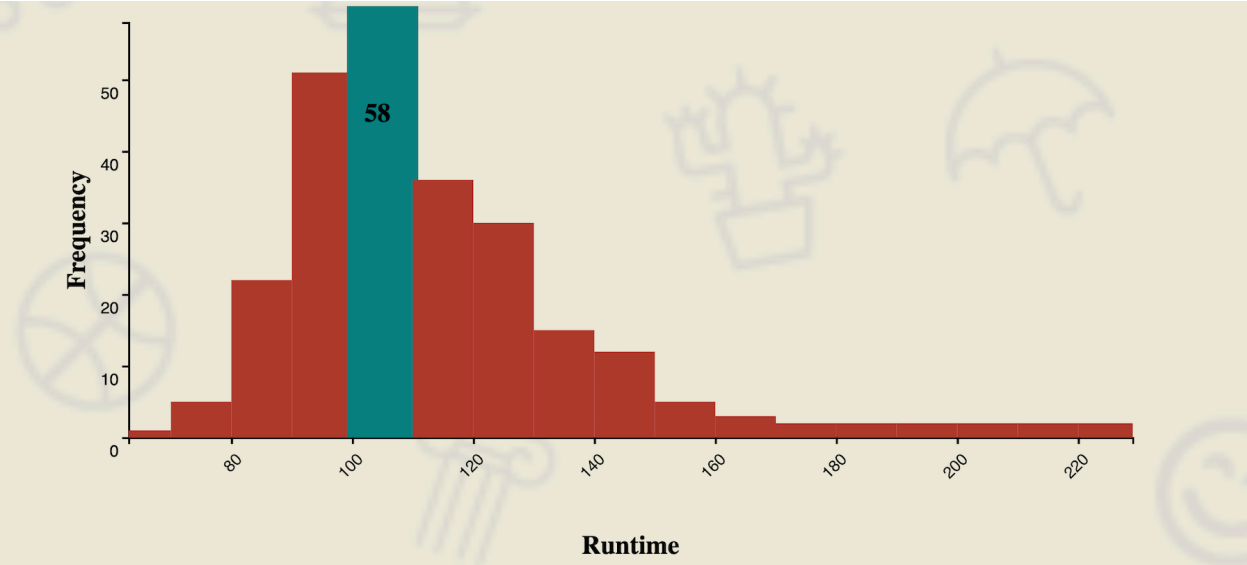
For each of 6 the numerical attributes(Runtime,Year,ImdbRating,Popularity ,Total votes and budget) ,I have created a histogram with the range on the x axis and its frequency on the y-axis.

Functionalities of histogram:

First functionality of the histogram is same as that of the bar chart.

1. On mouse-over 3 there are 3 functionalities:

   a. The value of the bar graph is displayed on the top of the cursor
   b. The bar becomes wider and higher
   c. The color of the bar changes.



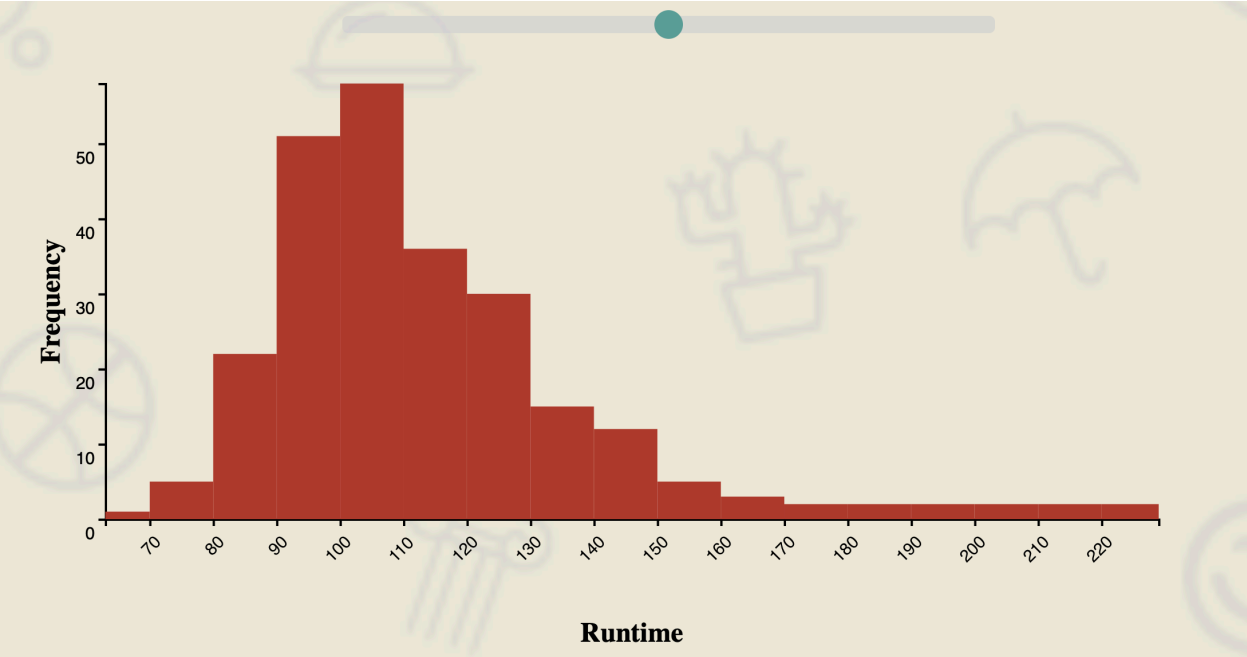
The implementation of this part is same as that of the bar graph with some minor changes of x and y positions.

2. On moving the slider to the right side the number of bins decrease and the bin width increases and on moving it to the left side the bin count increases and bin width decreases.

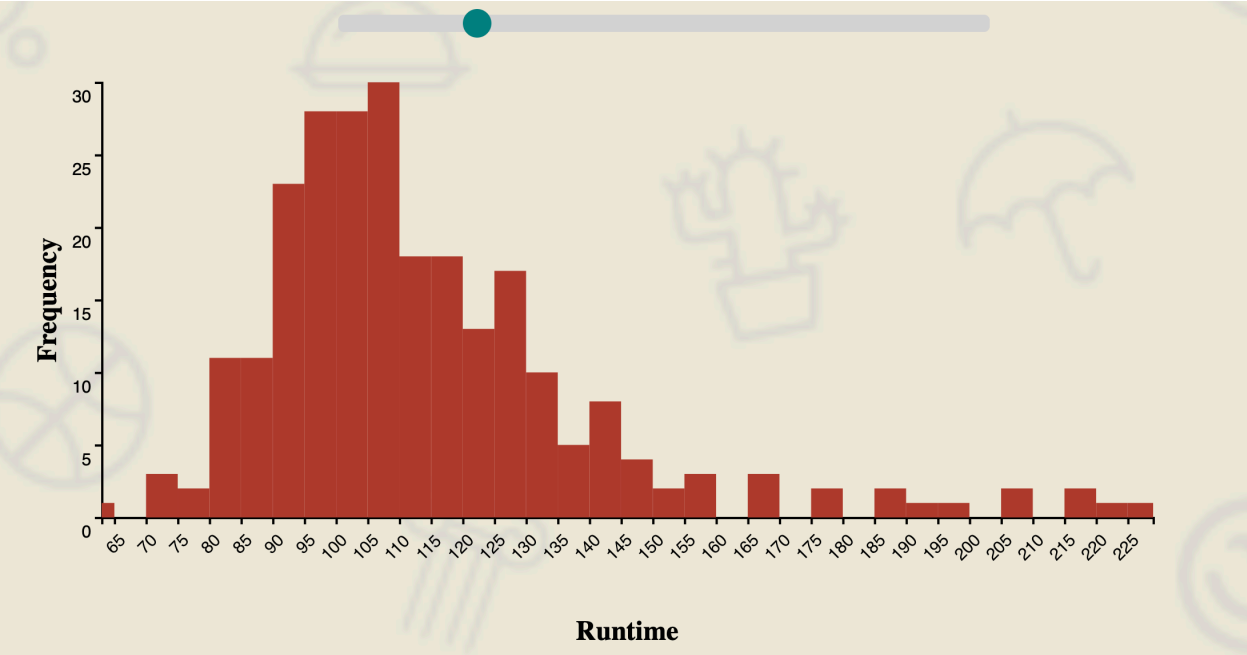
I have added a slider in this part in the html page which gets active only when a numerical value is selected .
The implementation of the slider is as follows:It takes the current value of the slider (sliderId.value) and calls the hist function again with the updated value of bins.

```
153
154    function changeBinWidth(map,bins,a){
155            d3.select("#slider").on("mousedown", function() {
156
157                var slider = document.getElementById("myRange");
158                slider.oninput = function() {
159                    d3.selectAll(".svg2").remove();
160                    var s=Math.abs(slider.value)
161                        hist(map,s,a)
162                }
163            })
164    }
```

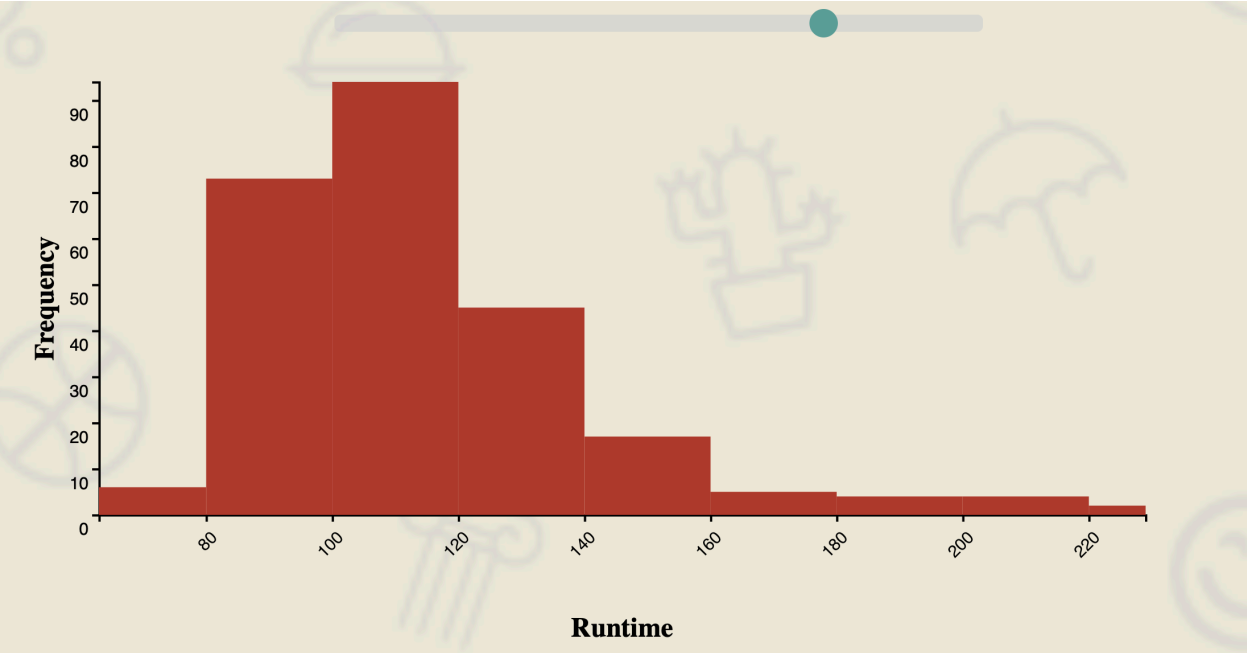
This figure shows the bar chart with the default bin count .



And as we move the slider towards left the number of bins increases and bin width decreases. Also the x-ticks increases with the bin numbers.



After moving the slider on the right ,bin count decreases, bin width increases and x-ticks decreases.



VIDEO LINK:

https://youtu.be/1gAc3xU6ulM