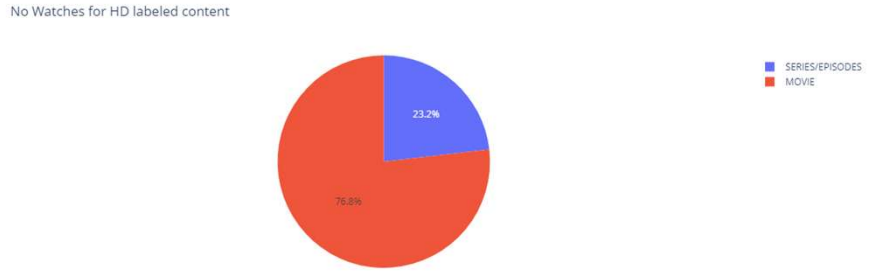
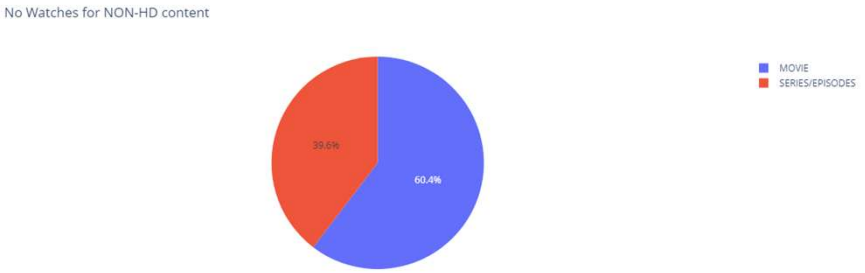
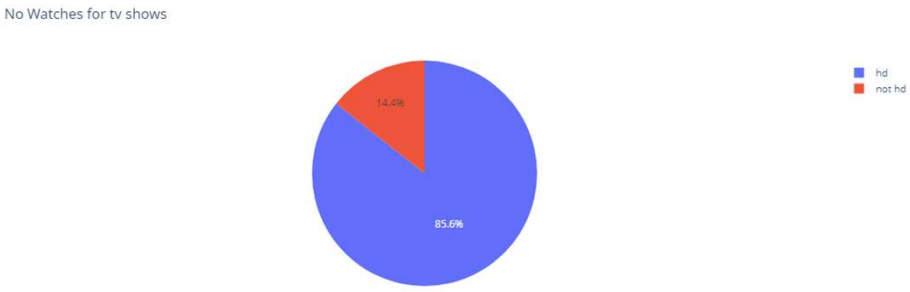
TASK#1
Rima Ali Alghamdi

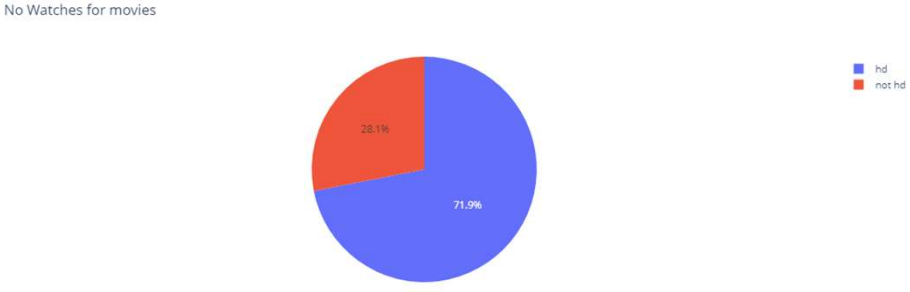
The following graphs compare between tv shows and movies in terms of number of watches, in the first figure all compared content was labeled with hd.
We can see that in terms of number of watches movies win, despite the quality of the video.





But does this mean that the quality of the content does not affect the numbers of watches? To answer this we must compare the number of view based on the quality. Each of the following graphs shows the variation of number of views for HD and non-HD content for each (tv shows and movies), as we can see, when it comes to tv shows, people prefer high quality, but in movies, in some cases low quality can be tolerable.





Code

```
isHD_grouped=df.copy()

isHD_grouped = isHD_grouped.groupby(['program_class', 'hd'])\
.agg({'user_id_maped': [('co1', 'nunique'),('co2', 'count')],\
      'duration_seconds': [('co3', 'sum')], }).reset_index()

isHD_grouped.columns = ['program_class','No of Users who Watched', 'No of
watches', 'Total watch time in seconds','hd']
isHD_grouped['Total watch time in houres']=isHD_grouped['Total watch time in
seconds']/3600
isHD_grouped = isHD_grouped.drop(columns=['Total watch time in seconds'])
isHD_grouped = isHD_grouped.drop(columns=['Total watch time in houres'])
isHD_grouped = isHD_grouped.drop(columns=['No of Users who Watched'])
isHD_grouped = isHD_grouped.sort_values(by=['hd'],
ascending=False).reset_index(drop=True)
isNotHD_grouped = isHD_grouped.tail(2)

isHD_grouped = isHD_grouped.head(2)
isHD_grouped

isNotHD_grouped

fig3 = px.pie(isNotHD_grouped, values='hd', names='program_class',\
             hover_data=['program_class'],title='No Watches for NON-HD content')
fig3.update_traces(sort=False)
fig3.show()

fig4 = px.pie(isHD_grouped, values='No of watches', names='program_class',\
             hover_data=['program_class'],title='No Watches for HD labeled
content')
fig4.update_traces(sort=False)
fig4.show()

tvshows = pd.DataFrame({'hd': ["hd","not hd"],
                    'No watches' : [isHD_grouped['hd'].values[0] ,
isNotHD_grouped['hd'].values[0]]
                    })
tvshows


movies = pd.DataFrame({'hd': ["hd","not hd"],
                    'No watches' : [isHD_grouped['hd'].values[1] ,
isNotHD_grouped['hd'].values[1]]
                    })
movies


fig5 = px.pie(tvshows, values='No watches', names='hd',\
             hover_data=['hd'],title='No Watches for tv shows')
fig5.update_traces(sort=False)
fig5.show()


fig6 = px.pie(movies, values='No watches', names='hd',\
             hover_data=['hd'],title='No Watches for movies')
fig6.update_traces(sort=False)
fig6.show()


content = tvshows.add(movies, fill_value=0)
content
```