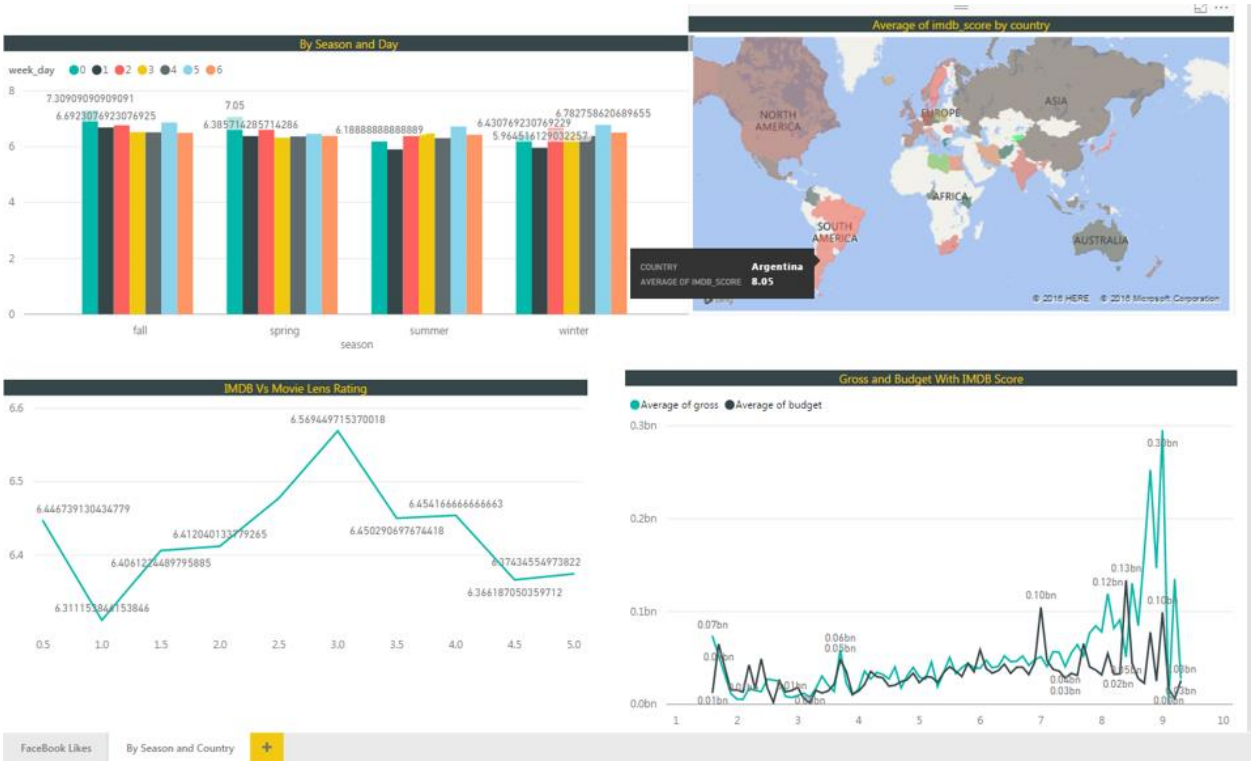
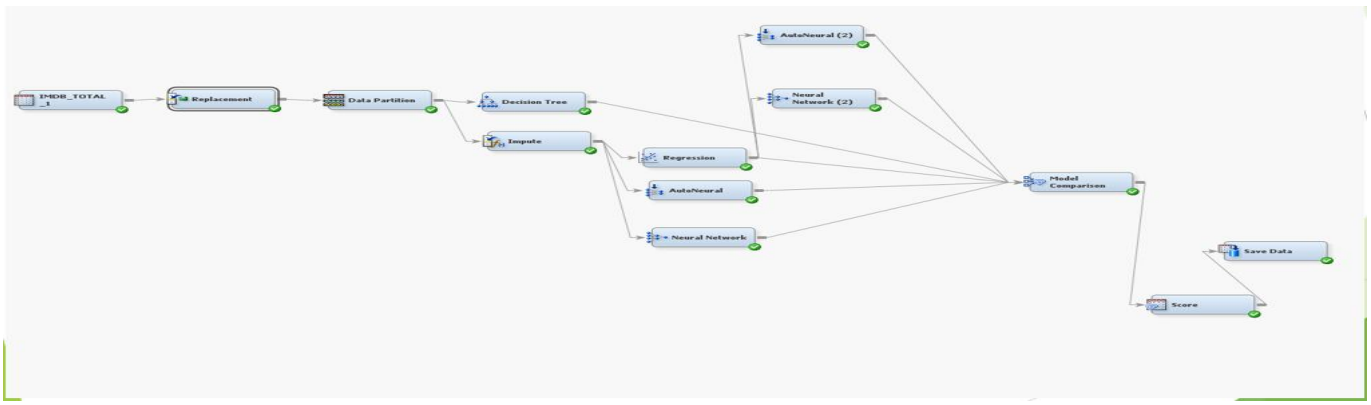
## ***Data Modeling Using PowerBI and SAS:***

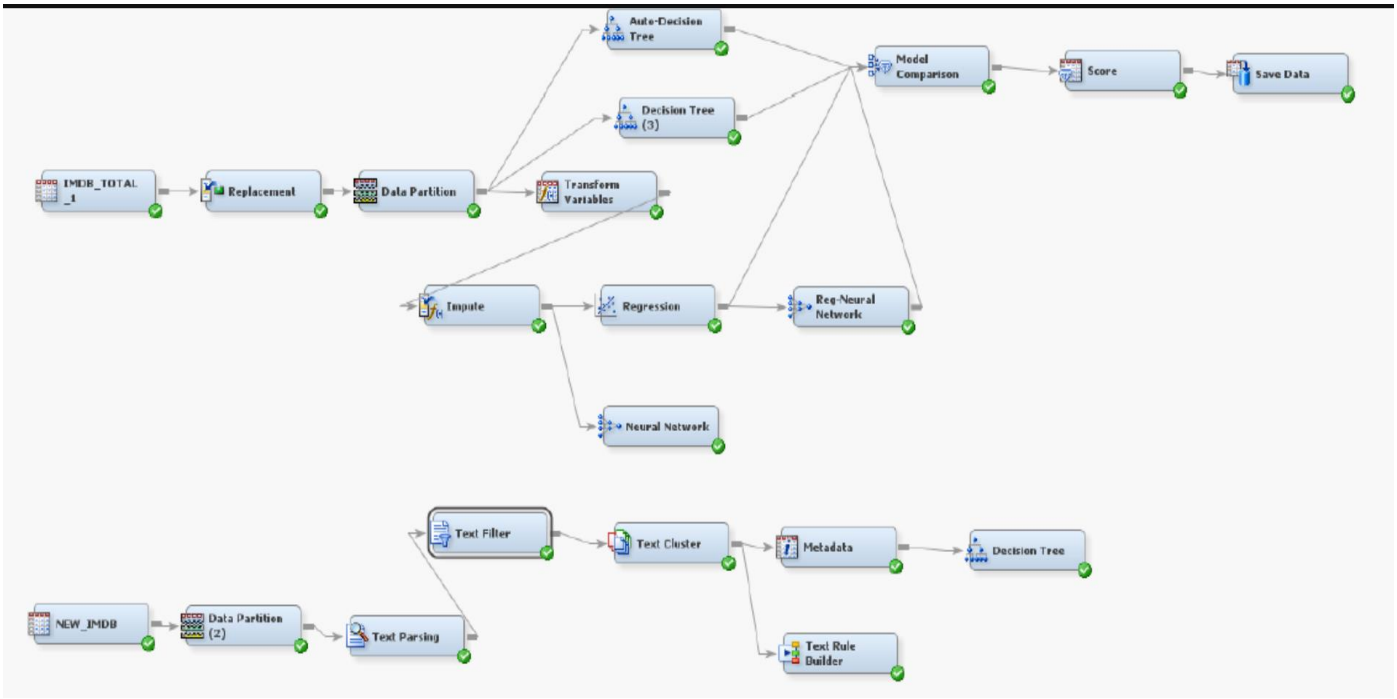


| Name | Role | Level | Report | Order | Drop | Lower Limit | Upper Limit |
|---|---|---|---|---|---|---|---|
| actor_1_facebook_likes | Input | Interval | No | | No | . | . |
| actor_1_name | Input | Nominal | No | | No | . | . |
| actor_2_facebook_likes | Rejected | Interval | No | | No | . | . |
| actor_2_name | Rejected | Nominal | No | | No | . | . |
| actor_3_facebook_likes | Rejected | Interval | No | | No | . | . |
| actor_3_name | Rejected | Nominal | No | | No | . | . |
| aspect_ratio | Input | Interval | No | | No | . | . |
| budget | Input | Interval | No | | No | . | . |
| cast_total_facebook_likes | Input | Interval | No | | No | . | . |
| color | Input | Binary | No | | No | . | . |
| country | Input | Nominal | No | | No | . | . |
| date | Input | Nominal | No | | No | . | . |
| director_facebook_likes | Input | Interval | No | | No | . | . |
| director_name | Input | Nominal | No | | No | . | . |
| duration | Input | Interval | No | | No | . | . |
| facenumber_in_poster | Input | Interval | No | | No | . | . |
| genres | Text | Nominal | No | | No | . | . |
| gross | Input | Interval | No | | No | . | . |
| imdb_score | Target | Interval | No | | No | . | . |
| language | Input | Nominal | No | | No | . | . |
| movie_facebook_likes | Input | Interval | No | | No | . | . |
| movie_imdb_link | Rejected | Nominal | No | | No | . | . |
| movie_lens_rating | Input | Interval | No | | No | . | . |
| movie_title | Text | Nominal | No | | No | . | . |
| num_critic_for_reviews | Input | Interval | No | | No | . | . |
| num_user_for_reviews | Input | Interval | No | | No | . | . |
| num_voted_users | Input | Interval | No | | No | . | . |
| plot_keywords | Input | Nominal | No | | No | . | . |
| season | Input | Nominal | No | | No | . | . |
| title_year | Input | Interval | No | | No | . | . |
| week_day | Input | Nominal | No | | No | . | . |

***Basic Model***

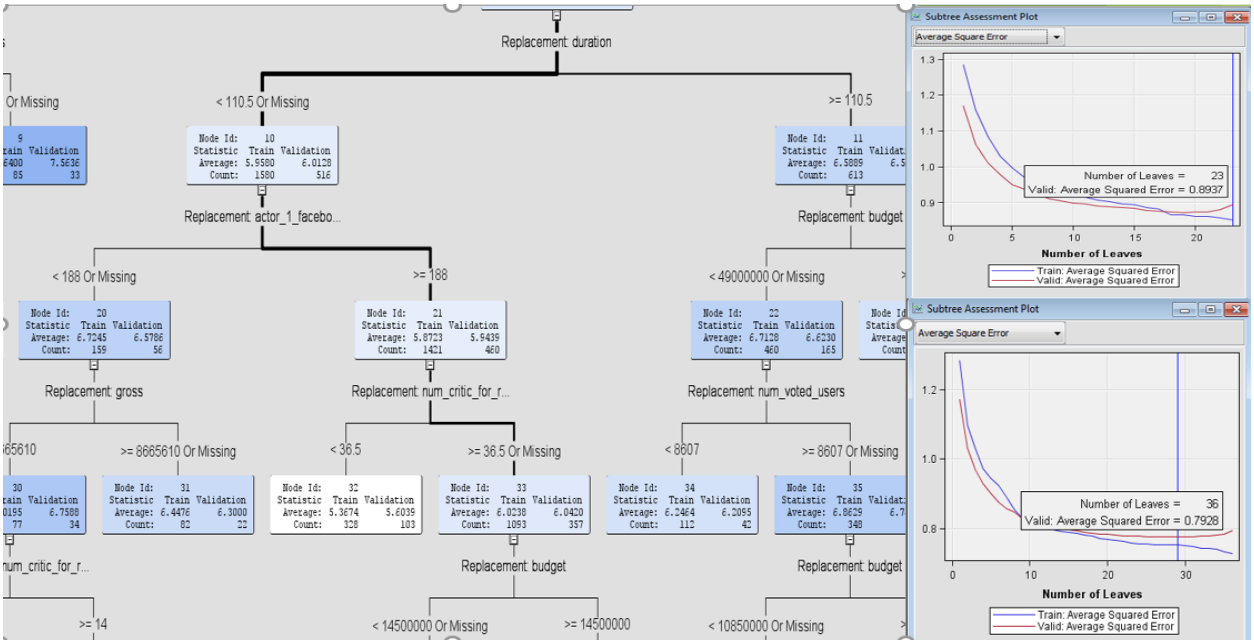


***Final Model:***




Analysis:

Decision Tree:



Model Comparison:

| Selected Model | Model Node | Model Description | Valid: Average Squared Error | Train: Average Squared Error |
|---|---|---|---|---|
| Y | Tree | Decision Tree | 0.77568 | 0.75261 |
| | AutoNeural2 | AutoNeural (2) | 1.13054 | 0.02750 |
| | Neural2 | Neural Network (2) | 1.13095 | 0.00340 |
| | Neural | Neural Network | 1.14057 | 0.56334 |
| | AutoNeural | AutoNeural | 1.14210 | 0.62261 |
| | Reg | Regression | 1.17545 | 0.00000 |

Final Predictions Vs Testing Data:

| movie_title | imdb_score | EM_PREDICTION |
|---|---|---|
| A Bugs Life | 7.2 | 6.524418605 |
| Amidst the Devils Wings | 4.3 | 6.307827789 |
| Big Mommas House 2 | 4.6 | 5.774398625 |
| Bram Stokers Dracula | 7.5 | 7.26 |
| Buffalo 66 | 7.5 | 6.307827789 |
| But Im a Cheerleader | 6.6 | 6.307827789 |
| Dancin Its On | 2.8 | 5.367378049 |
| Dave Chappelles Block Party | 7.2 | 6.307827789 |
| Devils Due | 4 | 6.307827789 |
| Dude, Wheres My Dog?! | 3.2 | 6.447560976 |
| Enders Game | 6.7 | 6.524418605 |
| Freddys Dead: The Final Nightmare | 4.9 | 6.307827789 |
| Gentlemans Agreement | 7.4 | 7.64 |
| Heavens Gate | 6.8 | 6.677826087 |
| I Dont Know How She Does It | 4.9 | 5.774398625 |
| Its a Mad, Mad, Mad, Mad World | 7.6 | 7.64 |
| Jesus Son | 7 | 6.307827789 |
| Kings Ransom | 4.1 | 5.367378049 |
| Lauberge espagnole | 7.3 | 7.223728814 |
| Lets Kill Wards Wife | 5.4 | 5.367378049 |
| Logans Run | 6.8 | 7.64 |
| Meeks Cutoff | 6.5 | 6.307827789 |
| Mr. Hollands Opus | 7.3 | 7.223728814 |
| My Best Friends Girl | 5.9 | 6.677826087 |
| My Bosss Daughter | 4.6 | 6.307827789 |
| Nims Island | 6 | 5.774398625 |
| On Her Majestys Secret Service | 6.8 | 7.64 |
| One Mans Hero | 6.2 | 6.246428571 |
| Schindlers List | 8.9 | 8.519230769 |
| Shes All That | 5.8 | 6.307827789 |