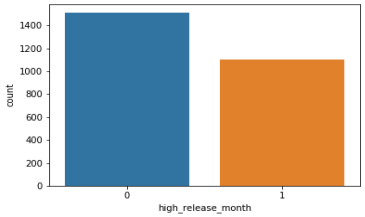
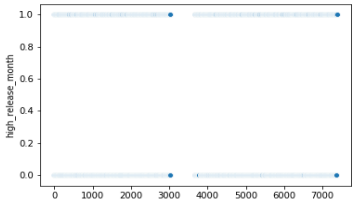
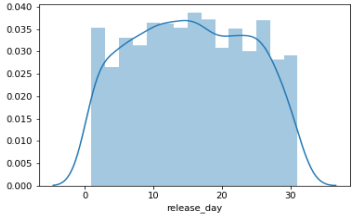
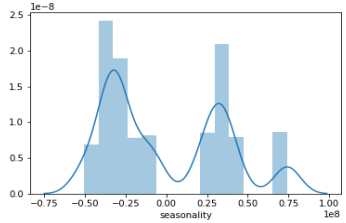
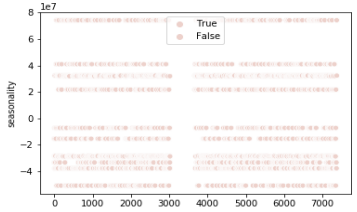
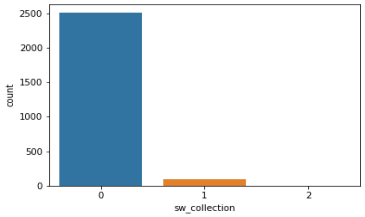
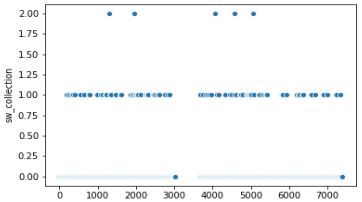
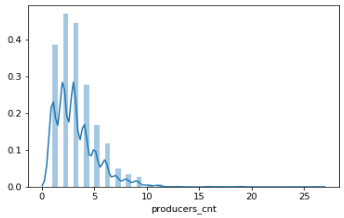
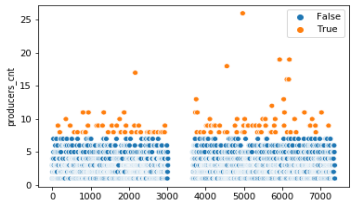
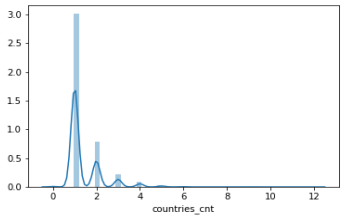
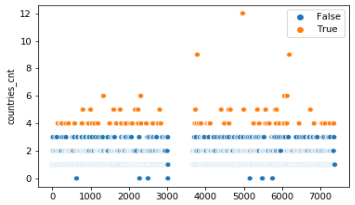
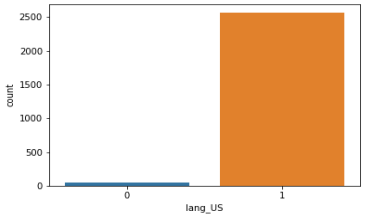
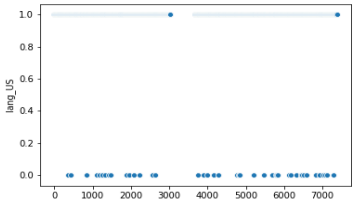
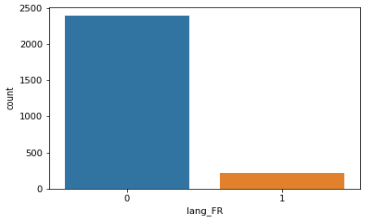
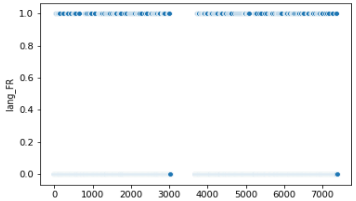
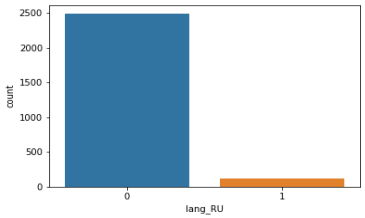
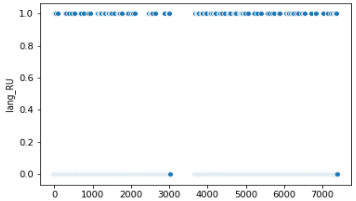
Exploratory Data Analysis (EDA)

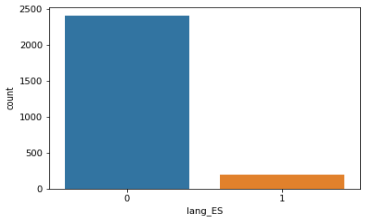
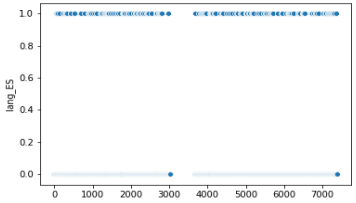
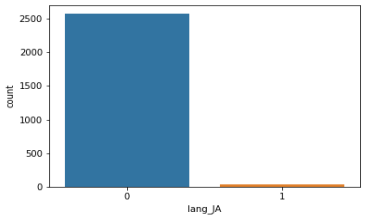
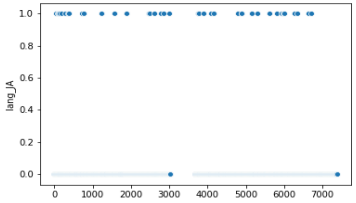
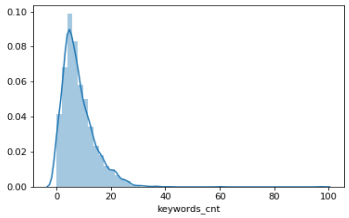
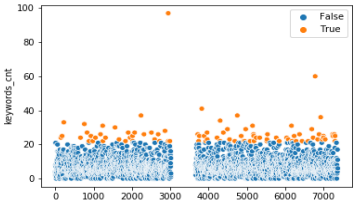
Show as …

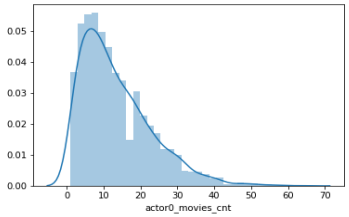
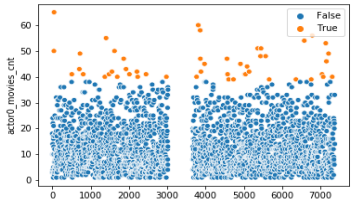
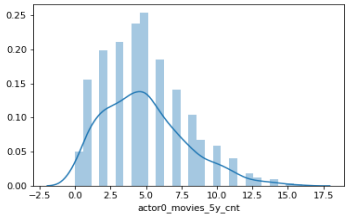
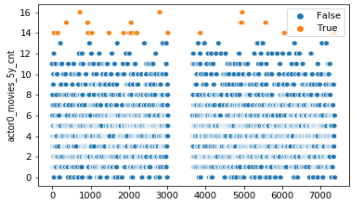
Variable	Distribution	Descriptive Statistics	Outliers	Dependent Variable Distribution
high_release_month	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 1,511.00 (57.80%) 1: 1,101.00 (42.20%)			
release_day	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%)		 Number of outliers: 0	

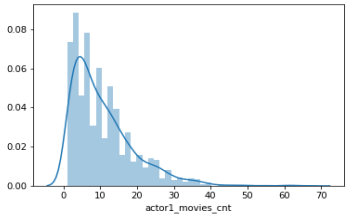
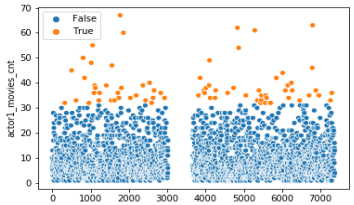
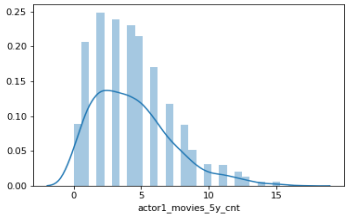
	Mean: 15.40 StdDev: 8.50 Median: 15.00 IQR: 8.00- 22.00 Min: 1.0 Max: 31.0 Kurtosis: 0.00 Skweness: -1.10		
seasonality	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: -1,047,664.50 StdDev: 37,002,138.40 Median: -15,309,317.00 IQR: -33,273,015.00- 32,322,910.00 Min: -50746591.0 Max: 74205073.0 Kurtosis: 0.50 Skweness: -1.00		 Number of outliers: 0
sw_collection	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%)		

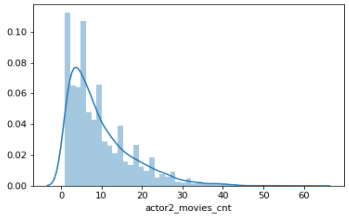
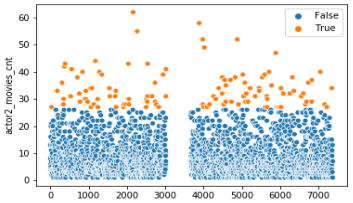
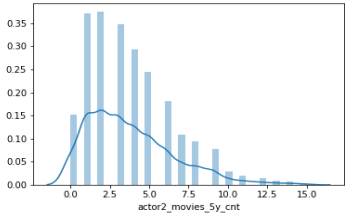
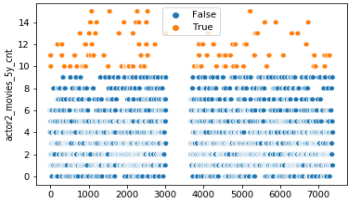
	Categories: 0: 2,509.00 (96.10%) 1: 98.00 (3.80%) 2: 5.00 (0.20%)		
producers_cnt	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 3.20 StdDev: 2.10 Median: 3.00 IQR: 2.00- 4.00 Min: 1.0 Max: 26.0 Kurtosis: 2.10 Skweness: 10.50		 Number of outliers: 110
countries_cnt	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 1.40 StdDev: 0.80 Median: 1.00 IQR: 1.00- 2.00 Min: 0 Max: 12		 Number of outliers: 81

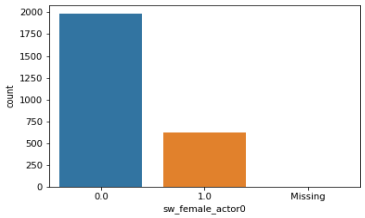
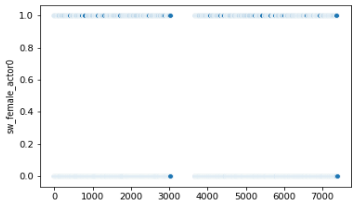
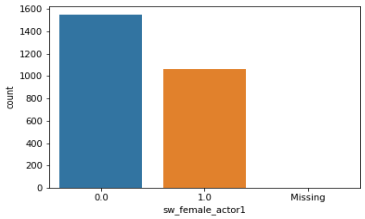
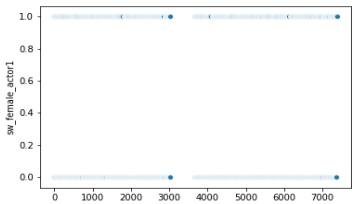
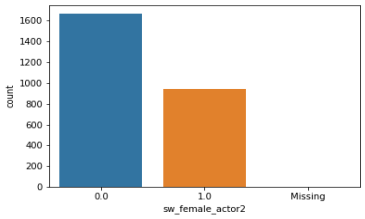
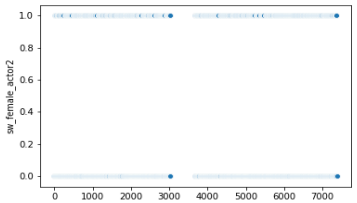
	Kurtosis: 3.30 Skweness: 20.10		
lang_US	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 1: 46.00 (1.80%) 0: 2,566.00 (98.20%)		
lang_FR	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,394.00 (91.70%) 1: 218.00 (8.30%)		
lang_RU	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,489.00 (95.30%) 1: 123.00 (4.70%)		

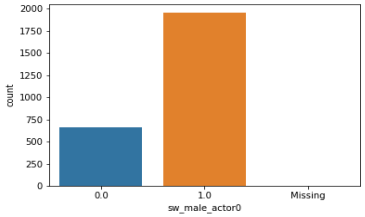
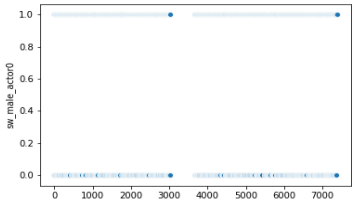
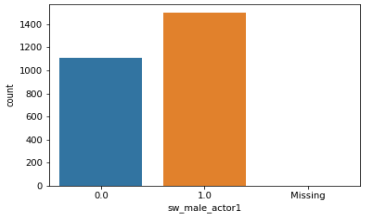
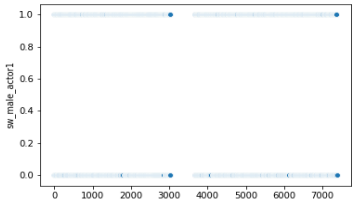
lang_ES	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,412.00 (92.30%) 1: 200.00 (7.70%)		
lang_JA	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,573.00 (98.50%) 1: 39.00 (1.50%)		
keywords_cnt	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 8.10 StdDev: 6.20 Median: 7.00 IQR: 4.00- 11.00 Min: 0 Max: 97		 Number of outliers: 90

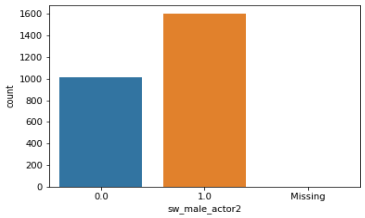
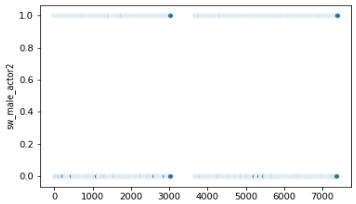
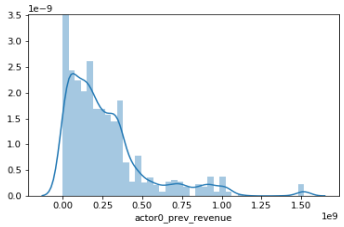
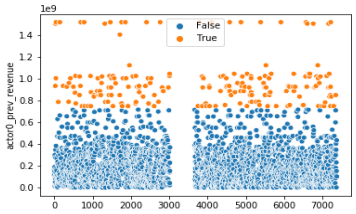
	Kurtosis: 2.40 Skweness: 19.40		
actor0_movies_cnt	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 13.40 StdDev: 9.60 Median: 11.00 IQR: 6.00- 19.00 Min: 1 Max: 65 Kurtosis: 1.20 Skweness: 1.60		 Number of outliers: 49
actor0_movies_5y_cnt	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 4.90 StdDev: 3.00 Median: 5.00 IQR: 3.00- 7.00 Min: 0 Max: 16 Kurtosis: 0.70 Skweness: 0.30		 Number of outliers: 23

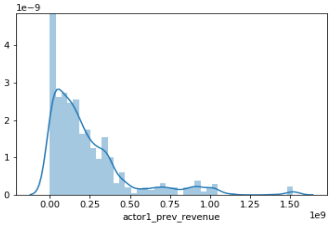
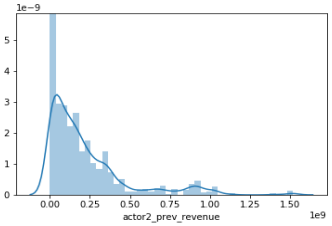
actor1_movies_cnt	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 10.80 StdDev: 8.60 Median: 9.00 IQR: 4.00- 15.00 Min: 1 Max: 67 Kurtosis: 1.60 Skweness: 3.90		 Number of outliers: 75
actor1_movies_5y_cnt	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 4.40 StdDev: 3.00 Median: 4.00 IQR: 2.00- 6.00 Min: 0 Max: 17 Kurtosis: 0.90 Skweness: 0.70		 Number of outliers: 38
actor2_movies_cnt	Data type:		

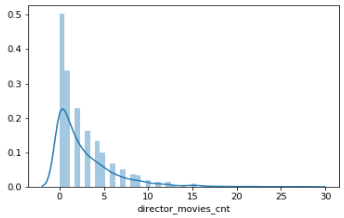
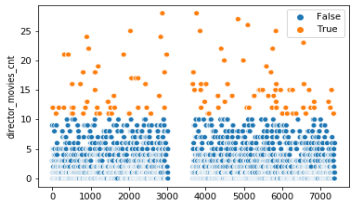
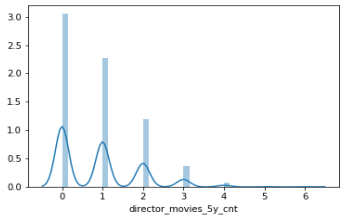
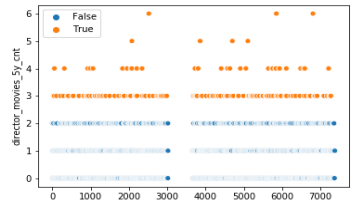
	Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 9.70 StdDev: 8.10 Median: 7.00 IQR: 4.00- 13.00 Min: 1 Max: 62 Kurtosis: 1.60 Skweness: 3.50		 Number of outliers: 116
actor2_movies_5y_cnt	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 3.80 StdDev: 2.80 Median: 3.00 IQR: 2.00- 5.00 Min: 0 Max: 15 Kurtosis: 1.00 Skweness: 1.00		 Number of outliers: 97
sw_female_actor0	Data type: Category		

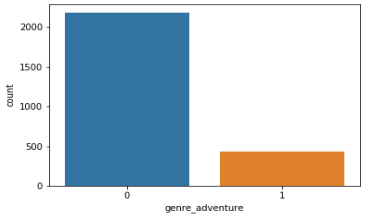
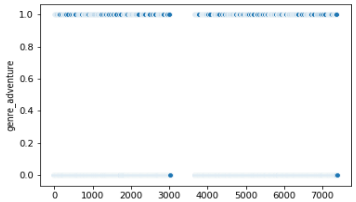
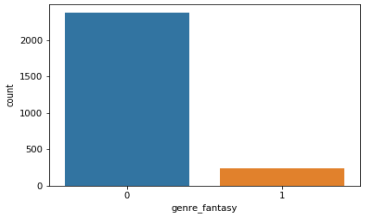
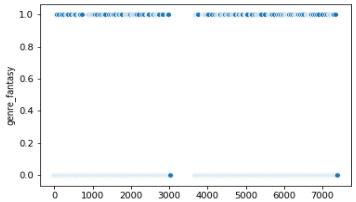
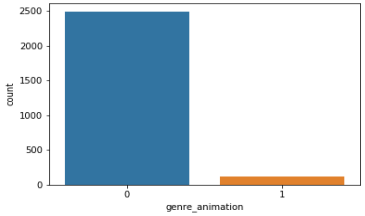
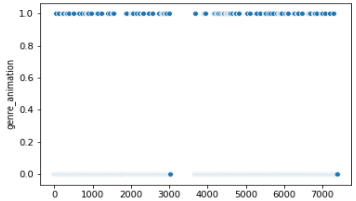
	Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0.0: 1,985.00 (76.00%) 1.0: 627.00 (24.00%) Missing: 0.00 (0.00%)		
sw_female_actor1	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0.0: 1,549.00 (59.30%) 1.0: 1,063.00 (40.70%) Missing: 0.00 (0.00%)		
sw_female_actor2	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0.0: 1,668.00 (63.90%)		

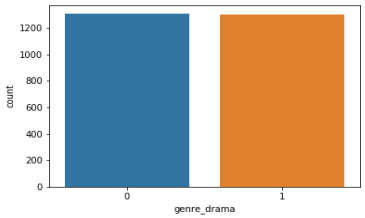
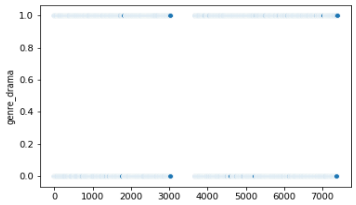
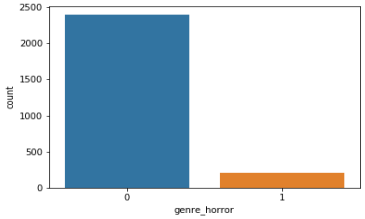
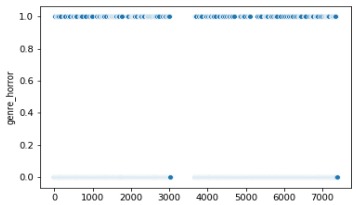
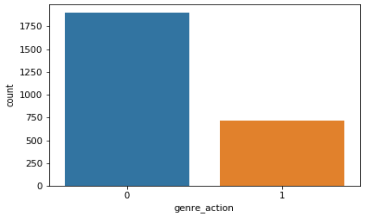
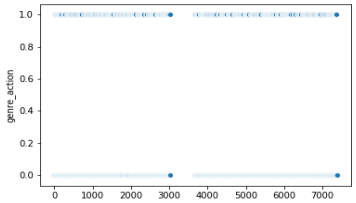
	1.0: 944.00 (36.10%) Missing: 0.00 (0.00%)		
sw_male_actor0	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 1.0: 658.00 (25.20%) 0.0: 1,954.00 (74.80%) Missing: 0.00 (0.00%)		
sw_male_actor1	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 1.0: 1,111.00 (42.50%) 0.0: 1,501.00 (57.50%) Missing: 0.00 (0.00%)		
sw_male_actor2	Data type: Category		

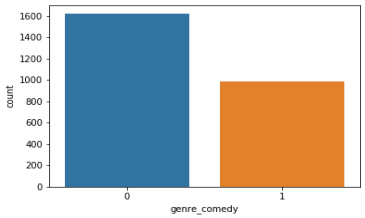
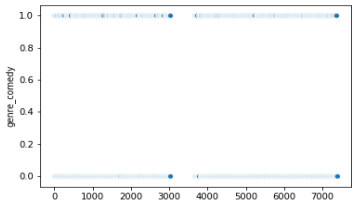
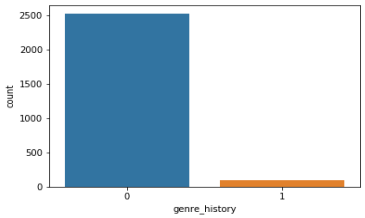
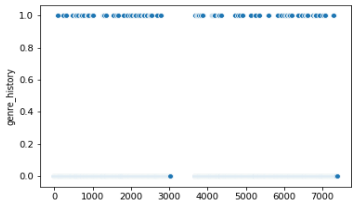
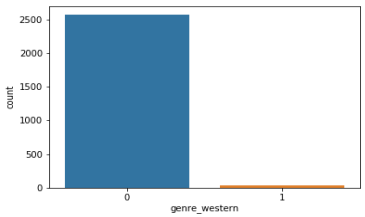
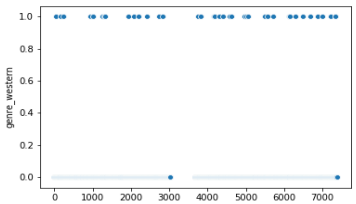
	Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 1.0: 1,010.00 (38.70%) 0.0: 1,602.00 (61.30%) Missing: 0.00 (0.00%)		
actor0_prev_revenue	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 269,397,280.70 StdDev: 263,168,760.50 Median: 196,482,882.00 IQR: 87,055,349.00-351,692,268.00 Min: 8.0 Max: 1519557910.0 Kurtosis: 1.90 Skweness: 4.50		 Number of outliers: 201
actor1_prev_revenue	Data type: Continuous		

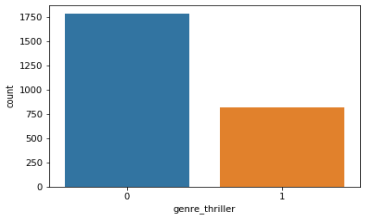
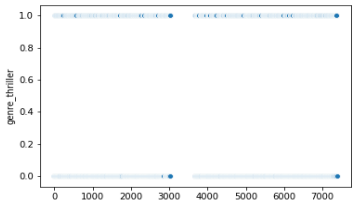
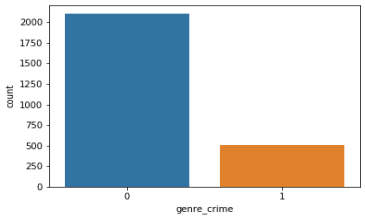
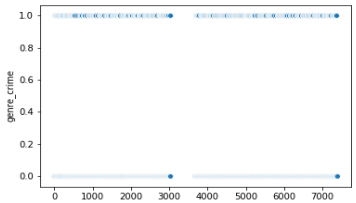
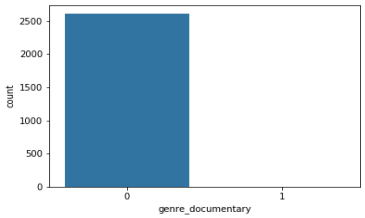
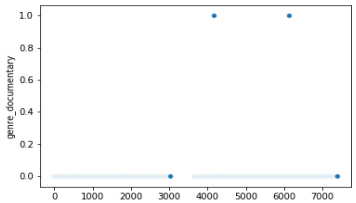
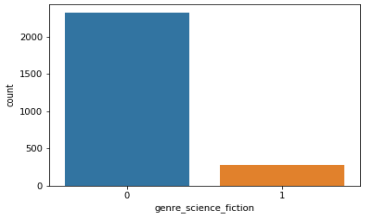
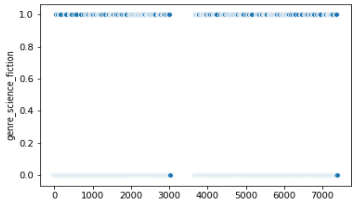
	Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 242,166,710.40 StdDev: 266,452,037.80 Median: 160,819,441.50 IQR: 63,104,668.80-313,542,341.00 Min: 79.0 Max: 1519557910.0 Kurtosis: 2.10 Skweness: 5.10		 Number of outliers: 224
actor2_prev_revenue	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 222,265,462.40 StdDev: 256,493,187.90 Median: 143,425,598.50 IQR: 43,313,102.20-294,456,605.00		 Number of outliers: 206

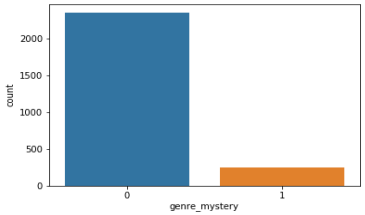
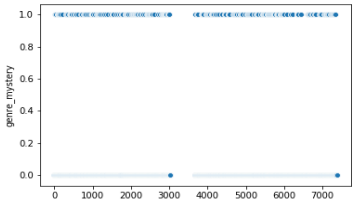
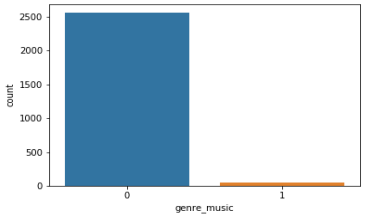
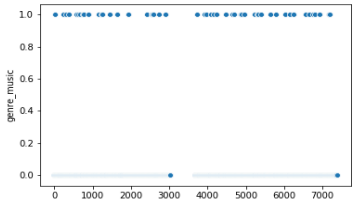
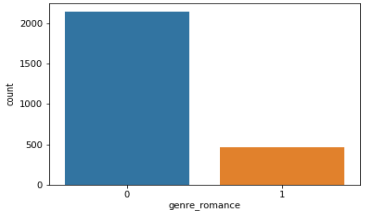
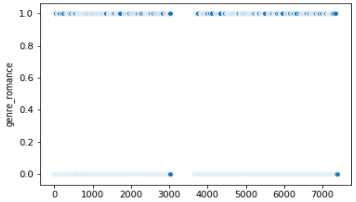
	Min: 1.0 Max: 1519557910.0 Kurtosis: 2.10 Skweness: 4.80		
director_movies_cnt	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 2.90 StdDev: 3.70 Median: 2.00 IQR: 0.00- 4.00 Min: 0 Max: 28 Kurtosis: 2.50 Skweness: 8.80		 Number of outliers: 114
director_movies_5y_cnt	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 0.90 StdDev: 1.00 Median: 1.00 IQR: 0.00- 1.00 Min: 0 Max: 6		 Number of outliers: 177

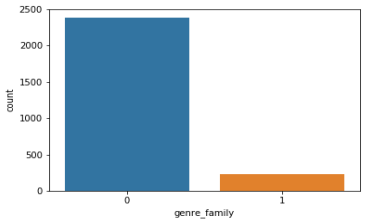
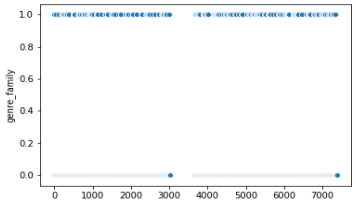
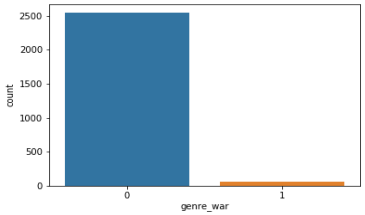
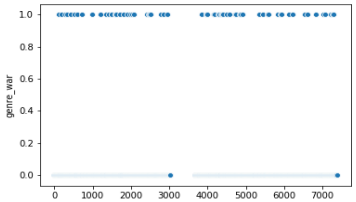
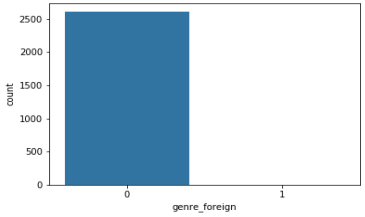
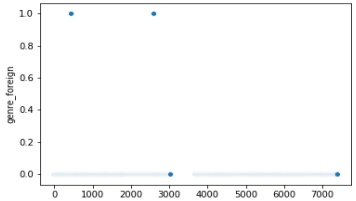
	Kurtosis: 1.10 Skweness: 1.20		
genre_adventure	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,184.00 (83.60%) 1: 428.00 (16.40%)		
genre_fantasy	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,375.00 (90.90%) 1: 237.00 (9.10%)		
genre_animation	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,488.00 (95.30%) 1: 124.00 (4.70%)		

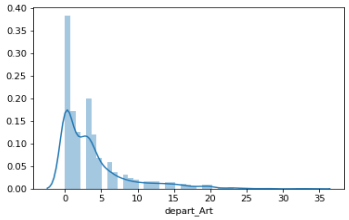
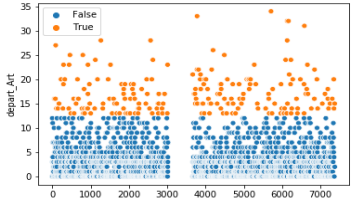
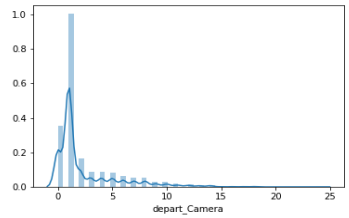
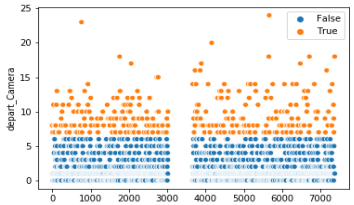
genre_drama	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 1,309.00 (50.10%) 1: 1,303.00 (49.90%)		
genre_horror	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,398.00 (91.80%) 1: 214.00 (8.20%)		
genre_action	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 1,900.00 (72.70%) 1: 712.00 (27.30%)		
	Data type: Category		

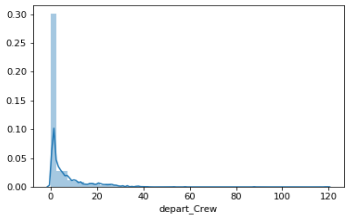
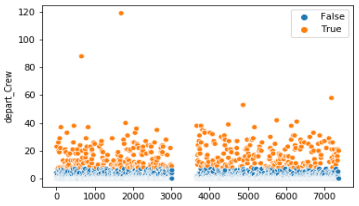
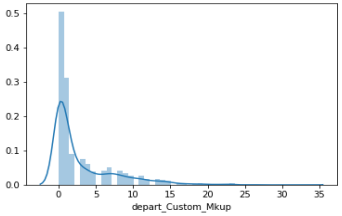
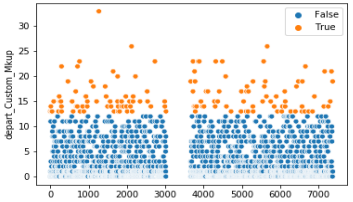
genre_comedy	Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 1,622.00 (62.10%) 1: 990.00 (37.90%)		
genre_history	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,520.00 (96.50%) 1: 92.00 (3.50%)		
genre_western	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,571.00 (98.40%) 1: 41.00 (1.60%)		
genre_thriller	Data type: Category Data length: 2612/2612		

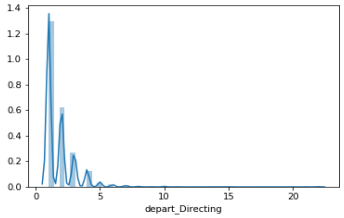
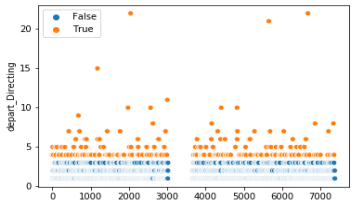
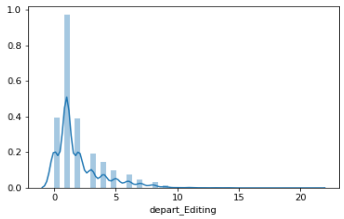
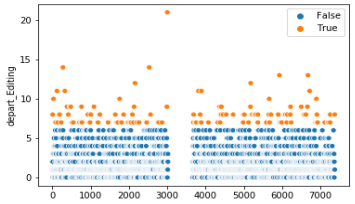
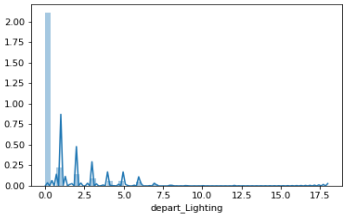
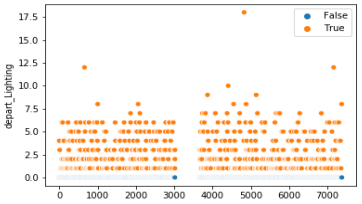
	Missing: 0 (0.00%%) Categories: 0: 1,788.00 (68.50%) 1: 824.00 (31.50%)		
genre_crime	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,103.00 (80.50%) 1: 509.00 (19.50%)		
genre_documentary	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,610.00 (99.90%) 1: 2.00 (0.10%)		
genre_science_fiction	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%)		

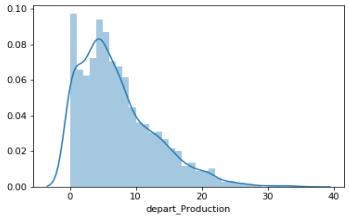
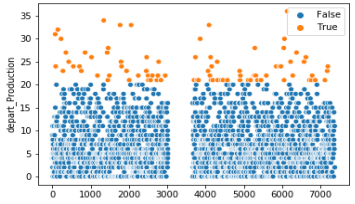
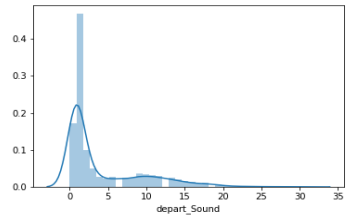
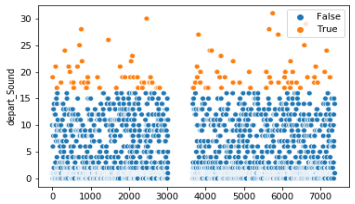
	Categories: 0: 2,326.00 (89.10%) 1: 286.00 (10.90%)		
genre_mystery	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,357.00 (90.20%) 1: 255.00 (9.80%)		
genre_music	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,560.00 (98.00%) 1: 52.00 (2.00%)		
genre_romance	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,143.00 (		

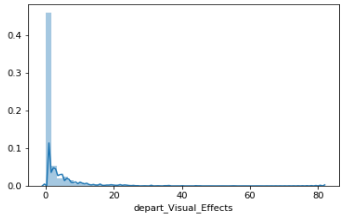
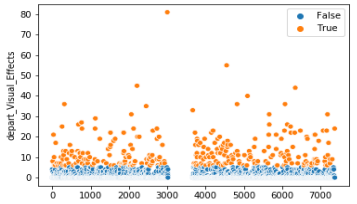
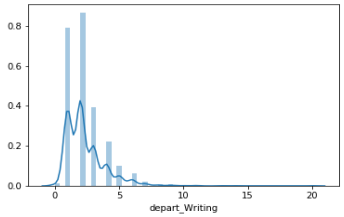
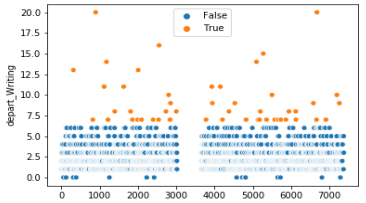
	82.00%) 1: 469.00 (18.00%)		
genre_family	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,386.00 (91.30%) 1: 226.00 (8.70%)		
genre_war	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,546.00 (97.50%) 1: 66.00 (2.50%)		
genre_foreign	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0: 2,610.00 (99.90%) 1: 2.00 (0.10%)		

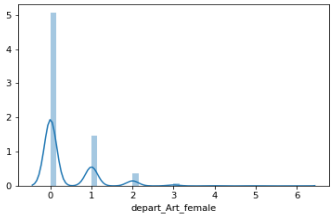
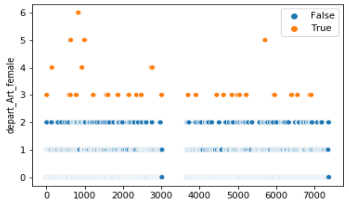
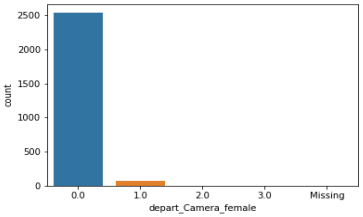
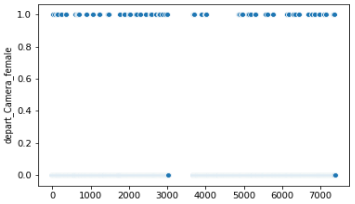
depart_Art	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 3.90 StdDev: 5.00 Median: 3.00 IQR: 0.00- 5.00 Min: 0.0 Max: 34.0 Kurtosis: 2.10 Skweness: 5.00		 Number of outliers: 207
depart_Camera	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 2.40 StdDev: 3.10 Median: 1.00 IQR: 1.00- 3.00 Min: 0.0 Max: 24.0 Kurtosis: 2.20 Skweness: 5.60		 Number of outliers: 296
depart_Crew	Data type: Continuous		

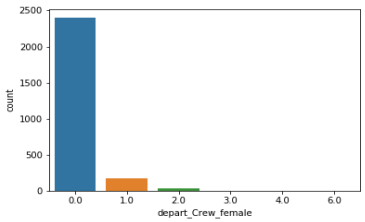
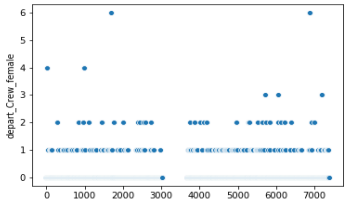
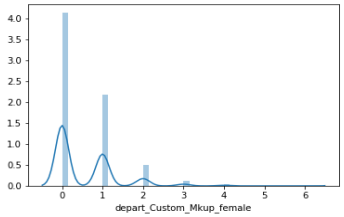
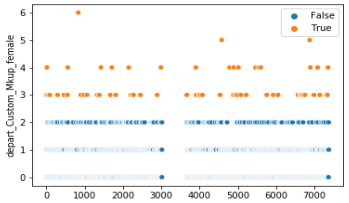
	Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 3.60 StdDev: 7.40 Median: 0.00 IQR: 0.00- 3.00 Min: 0.0 Max: 119.0 Kurtosis: 4.20 Skweness: 34.10		 Number of outliers: 398
depart_Custom_Mkup	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 3.10 StdDev: 4.40 Median: 1.00 IQR: 0.00- 5.00 Min: 0.0 Max: 33.0 Kurtosis: 1.90 Skweness: 4.00		 Number of outliers: 135
depart_Directing	Data type: Continuous Data length: 2612/2612 (		

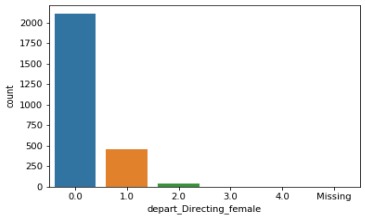
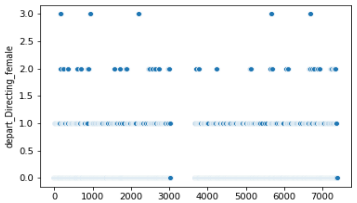
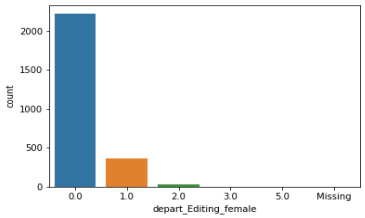
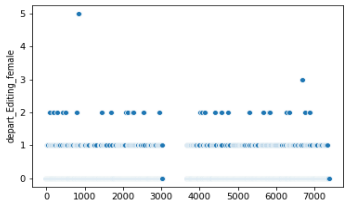
	100.00%%) Missing: 0 (0.00%%) Mean: 1.80 StdDev: 1.40 Median: 1.00 IQR: 1.00- 2.00 Min: 1.0 Max: 22.0 Kurtosis: 5.20 Skweness: 55.90		 Number of outliers: 220
depart_Editing	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 2.00 StdDev: 2.00 Median: 1.00 IQR: 1.00- 3.00 Min: 0.0 Max: 21.0 Kurtosis: 2.00 Skweness: 6.20		 Number of outliers: 118
depart_Lighting	Data type: Continuous Data length: 2612/2612 (100.00%%)		

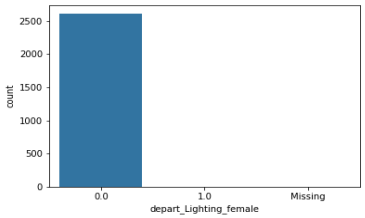
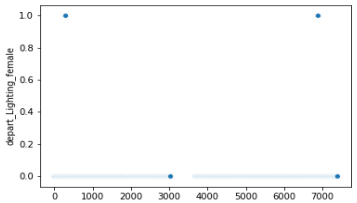
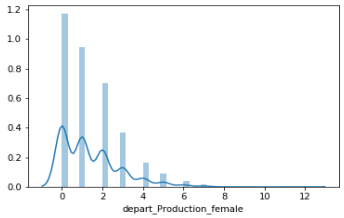
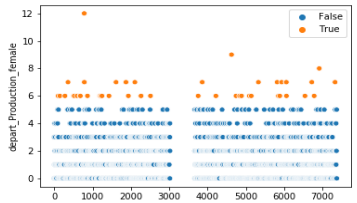
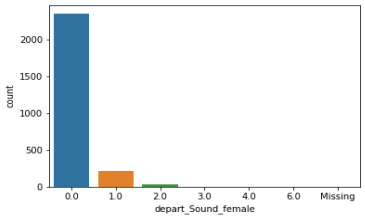
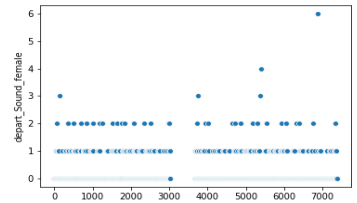
	Missing: 0 (0.00%%) Mean: 0.70 StdDev: 1.60 Median: 0.00 IQR: 0.00- 0.00 Min: 0.0 Max: 18.0 Kurtosis: 3.10 Skweness: 12.90		Number of outliers: 627
depart_Production	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 7.10 StdDev: 5.80 Median: 6.00 IQR: 3.00- 10.00 Min: 0.0 Max: 36.0 Kurtosis: 1.10 Skweness: 1.40		 Number of outliers: 85
depart_Sound	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%)		 Number of outliers: 97

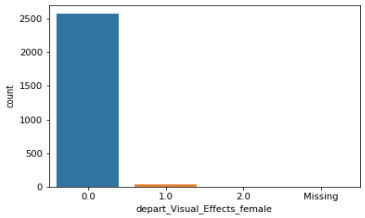
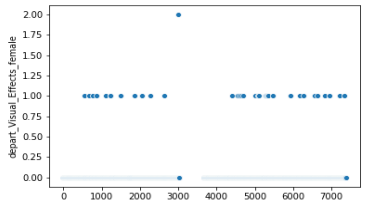
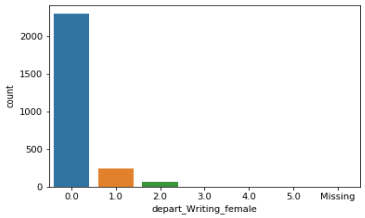
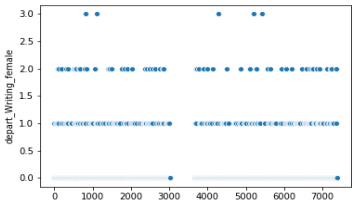
	Mean: 4.20 StdDev: 5.30 Median: 1.00 IQR: 1.00- 7.00 Min: 0.0 Max: 31.0 Kurtosis: 1.60 Skweness: 2.00		
depart_Visual_Effects	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 2.00 StdDev: 4.90 Median: 0.00 IQR: 0.00- 2.00 Min: 0.0 Max: 81.0 Kurtosis: 5.20 Skweness: 44.20		 Number of outliers: 290
depart_Writing	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 2.40 StdDev: 1.70		 Number of outliers: 54

	Median: 2.00 IQR: 1.00- 3.00 Min: 0.0 Max: 20.0 Kurtosis: 2.90 Skweness: 17.40		
depart_Art_female	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 0.40 StdDev: 0.70 Median: 0.00 IQR: 0.00- 1.00 Min: 0.0 Max: 6.0 Kurtosis: 2.30 Skweness: 7.30		 Number of outliers: 35
depart_Camera_female	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0.0: 2,536.00 (97.10%) 1.0: 76.00 (2.90%) Missing: 0.00 (0.00%)		

	3.0: 0.00 (0.00%) 2.0: 0.00 (0.00%)		
depart_Crew_female	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0.0: 2,401.00 (91.90%) 1.0: 174.00 (6.70%) 2.0: 30.00 (1.10%) 3.0: 3.00 (0.10%) 6.0: 2.00 (0.10%) 4.0: 2.00 (0.10%)		
depart_Custom_Mkup_female	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 0.50 StdDev: 0.80 Median: 0.00 IQR: 0.00- 1.00 Min: 0.0 Max: 6.0 Kurtosis: 1.80 Skweness: 4.20		 Number of outliers: 63
depart_Directing_female	Data type: Category		

	Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0.0: 2,109.00 (80.70%) 1.0: 457.00 (17.50%) 2.0: 41.00 (1.60%) 3.0: 5.00 (0.20%) Missing: 0.00 (0.00%) 4.0: 0.00 (0.00%)		
depart_Editing_female	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0.0: 2,222.00 (85.10%) 1.0: 360.00 (13.80%) 2.0: 28.00 (1.10%) 5.0: 1.00 (0.00%) 3.0: 1.00 (0.00%) Missing: 1.00 (0.00%)		
depart_Lighting_female	Data type: Category Data length: 2612/2612		

	Missing: 0 (0.00%%) Categories: 0.0: 2,610.00 (99.90%) 1.0: 2.00 (0.10%) Missing: 0.00 (0.00%)		
depart_Production_female	Data type: Continuous Data length: 2612/2612 (100.00%%) Missing: 0 (0.00%%) Mean: 1.40 StdDev: 1.50 Median: 1.00 IQR: 0.00- 2.00 Min: 0.0 Max: 12.0 Kurtosis: 1.30 Skweness: 2.30		 Number of outliers: 45
depart_Sound_female	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0.0: 2,359.00 (90.30%)		

	1.0: 215.00 (8.20%) 2.0: 33.00 (1.30%) 3.0: 3.00 (0.10%) 6.0: 1.00 (0.00%) 4.0: 1.00 (0.00%) Missing: 1.00 (0.00%)		
depart_Visual_Effects_female	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0.0: 2,576.00 (98.60%) 1.0: 35.00 (1.30%) 2.0: 1.00 (0.00%) Missing: 0.00 (0.00%)		
depart_Writing_female	Data type: Category Data length: 2612/2612 Missing: 0 (0.00%%) Categories: 0.0: 2,297.00 (87.90%) 1.0: 244.00 (9.30%) 2.0: 66.00 (2.50%) 3.0: 5.00 (0.20%) Missing: 0.00 (0.00%)		

	5.0: 0.00 (0.00%) 4.0: 0.00 (0.00%)		