**תרגיל מסכם – EDA**

1. revenue:

ערכים קיימים: 2,993

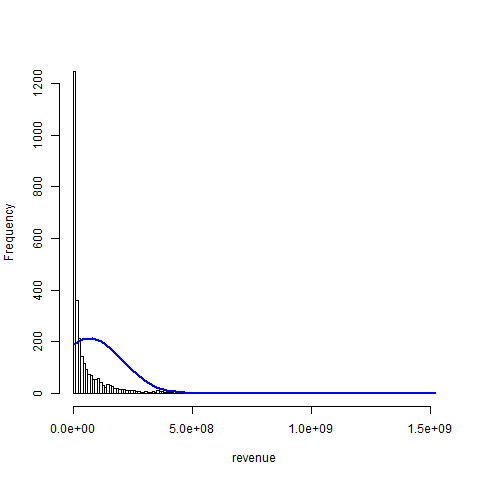
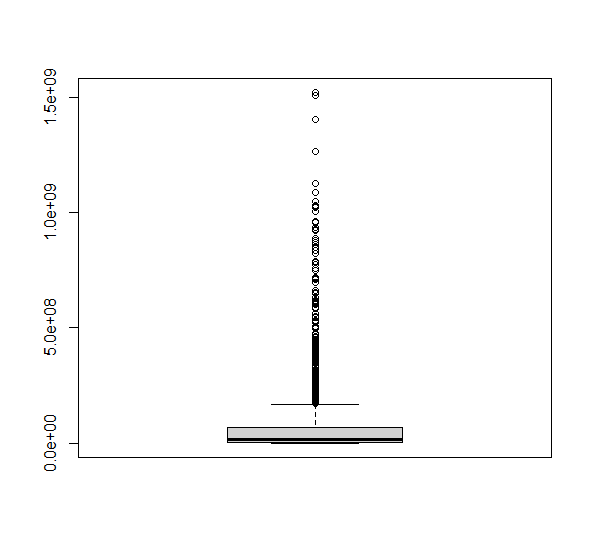
ערכים חסרים: 4,382

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 1 | 1.52E+09 | 6.69E+07 | 1.69E+07 | 8.80E+02 | 2.41E+06 | 6.91E+07 |

מספר ה- outliers: 415

תצוגה גרפית: לשנות טווח של X

שאלות פתוחות: כיצד להתייחס לחוסרים ולמספרים קטנים במיוחד?

1. Movie\_id

מפתח, ללא משמעות נוספת

ערכים קיימים: 7,375

ערכים חסרים: אין

1. movie\_title
2. budget
3. original\_language:

ערכים קיימים: 7,375

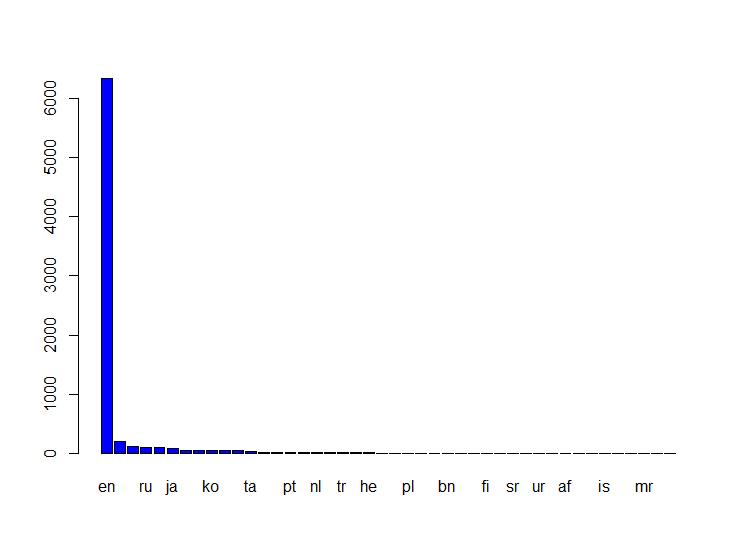
ערכים חסרים: אין

מספר קטגוריות: 44

השפה הפופולרית ביותר: En, עם 6336 תצפיות

שפות פופולריות: En, Fr, Hi, Ru, Es, Ja

תצוגה גרפית:



1. Popularity-

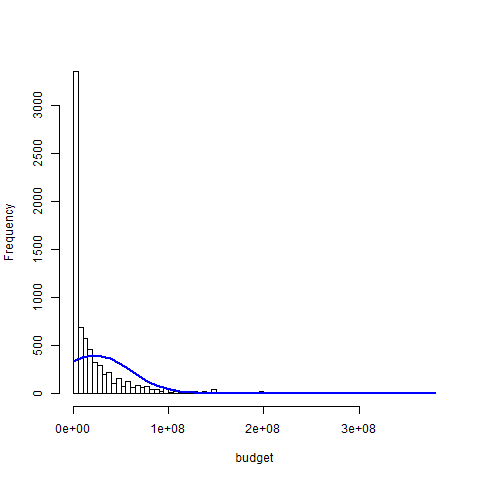
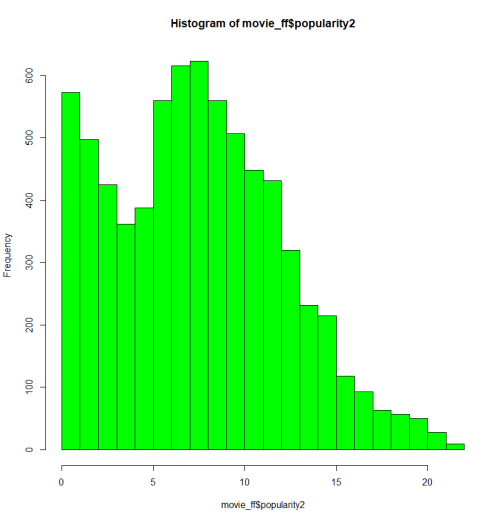
להסיר outliers

ערכים קיימים: 7,375

ערכים חסרים: אין

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 547.5 | 8.54 | 7.456 | 12.18 | 3.97 | 10.93 |
| בהגדרת 201 outliers כ- NA | | | | | | |
| 0 | 21.364 | 7.292 | 7.292 |  | 3.86 | 10.57 |

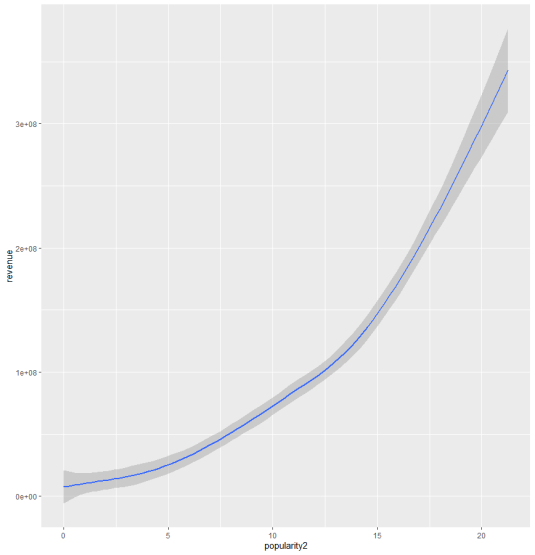
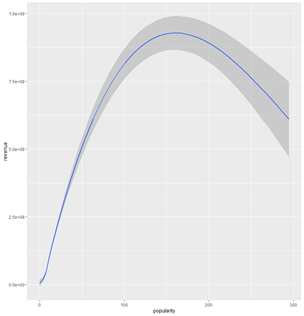
תצוגה גרפית - גולמי: without outliers



Spearman’s correlation

|  |  |  |  |
| --- | --- | --- | --- |
| rho =0.5945387 | S = 181183325 | p-value < 2.2e-16 | Row data |
| rho 0.5763747 | S = 176507326 | p-value < 2.2e-16 | Without outliers |

Row data –correlation : correlation without outliers



1. popularity\_group:

ערכים קיימים: 7,375

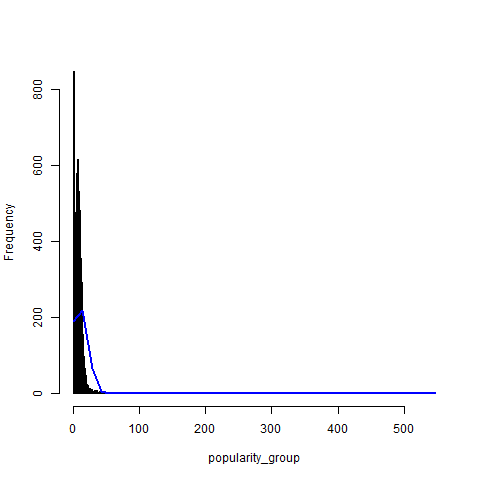
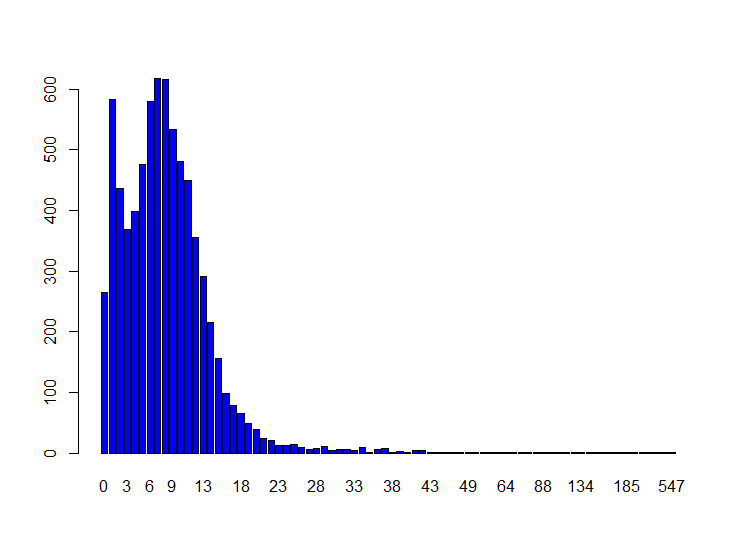
ערכים חסרים: אין

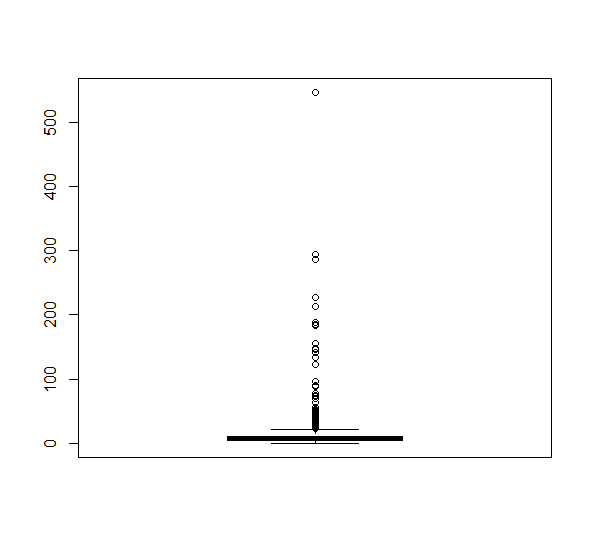
מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 547 | 8.532 | 7 | 12.18 | 4 | 11 |

מספר ה- outliers: 200

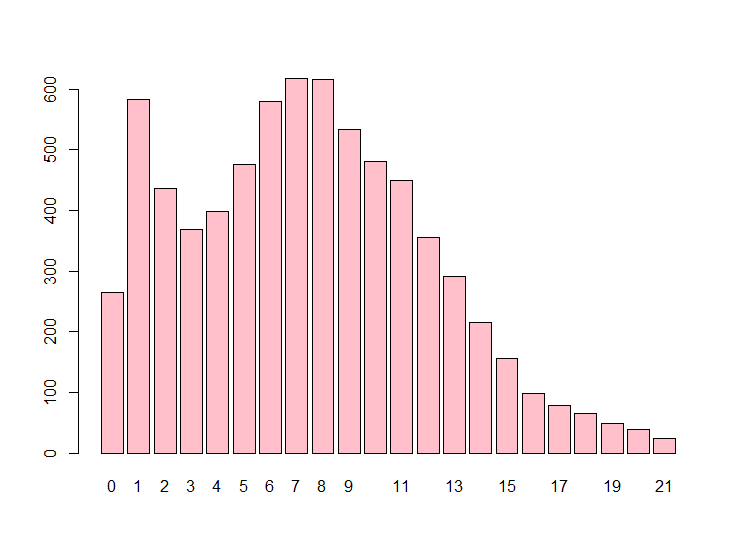
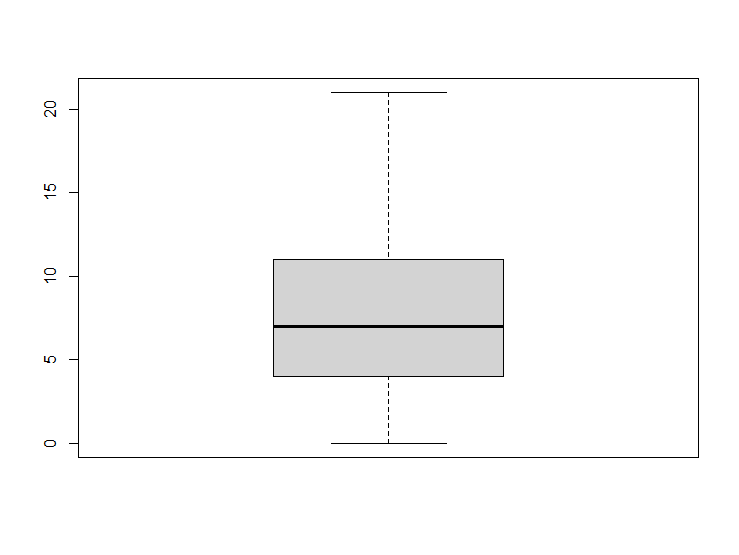
תצוגה גרפית:



ניתוח לאחר הוצאת ה- outliers:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 21 | 7.462 | 7 |  | 4 | 11 |

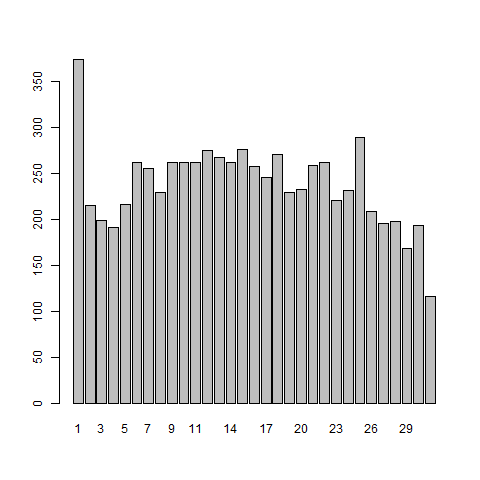
1. release\_date

ערכים קיימים:7374

ערכים חסרים:1

קטגוריות: 31

תצוגה גרפית



1. runtime:

ערכים קיימים: 7,370

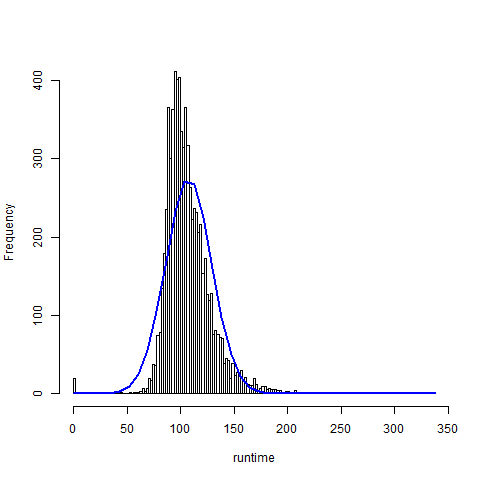
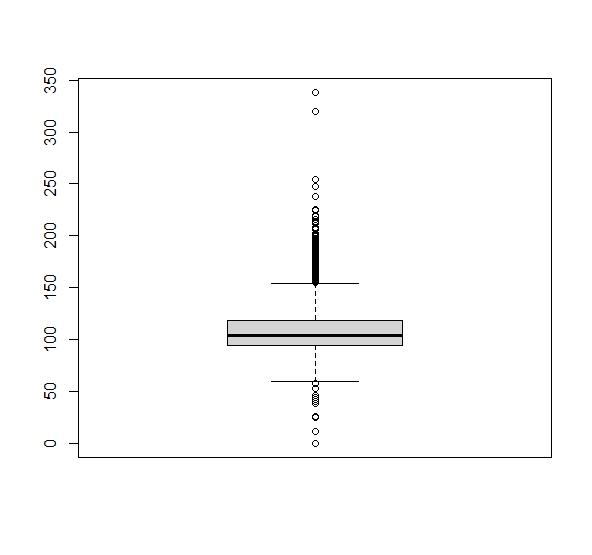
ערכים חסרים: 5

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 338 | 107.7 | 104 | 21.41 | 94 | 118 |

מספר ה- outliers: 275

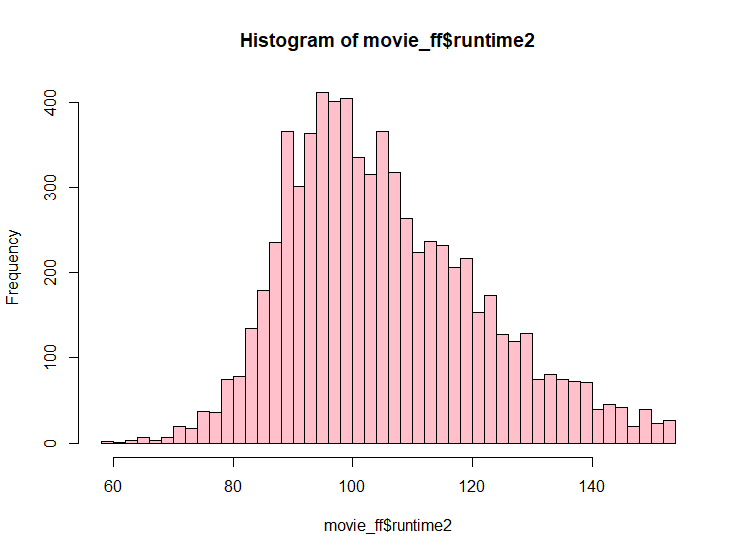
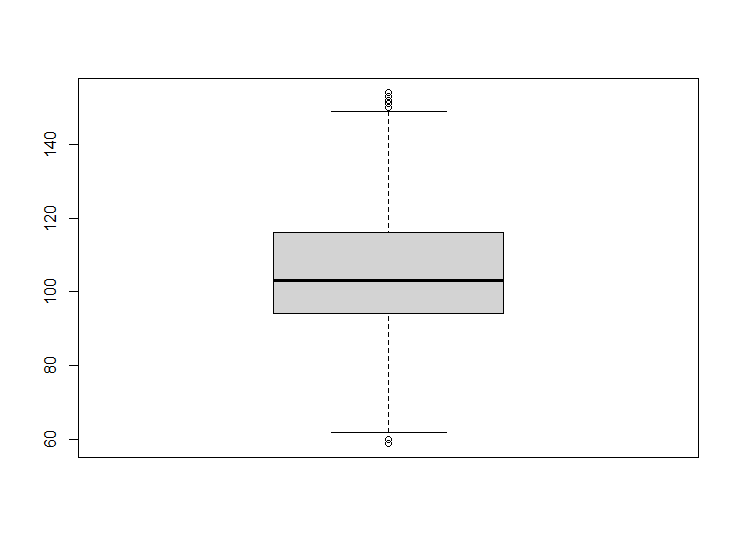
תצוגה גרפית:

חשוב: להתייחס ל- NULL ולסרטים קצרים במיוחד

ניתוח לאחר הוצאת ה- outliers:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 59 | 154 | 105.9 | 103 |  | 94 | 116 |

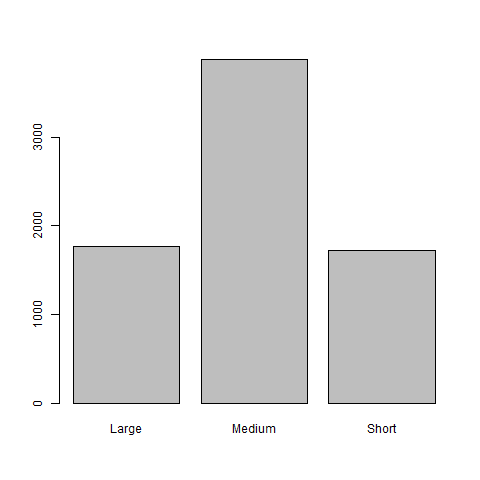
 

1. runtime\_cat

ערכים קיימים: 7,375

ערכים חסרים: אין

מספר קטגוריות: 3 (Large : 1769 ,Medium : 3884, Short : 1722)

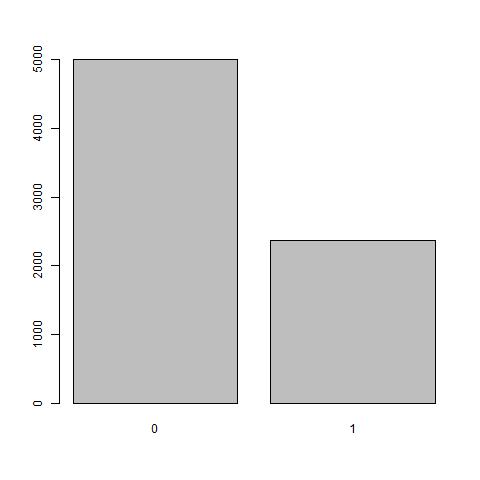


1. sw\_lang\_en
2. sw\_web\_presence:

ערכים קיימים: 7,375

ערכים חסרים: אין

מספר קטגוריות: 2 (5,012 – אין, 2,363 – יש)



1. sw\_has\_poster

ערכים קיימים: 7,375

ערכים חסרים: אין

קטגוריות :2 (0- 2, 1- 7373)



1. sw\_tagline
2. keyword\_cnt:

ערכים קיימים: 7,370

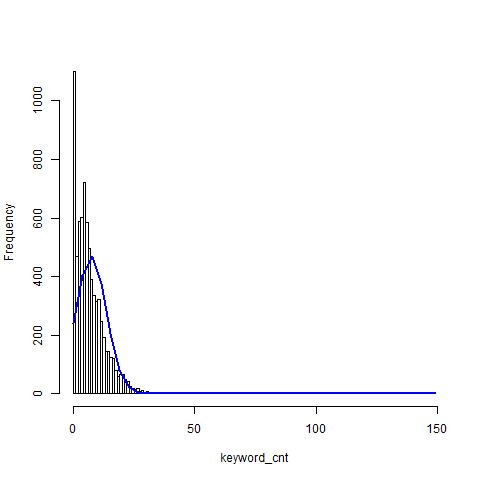
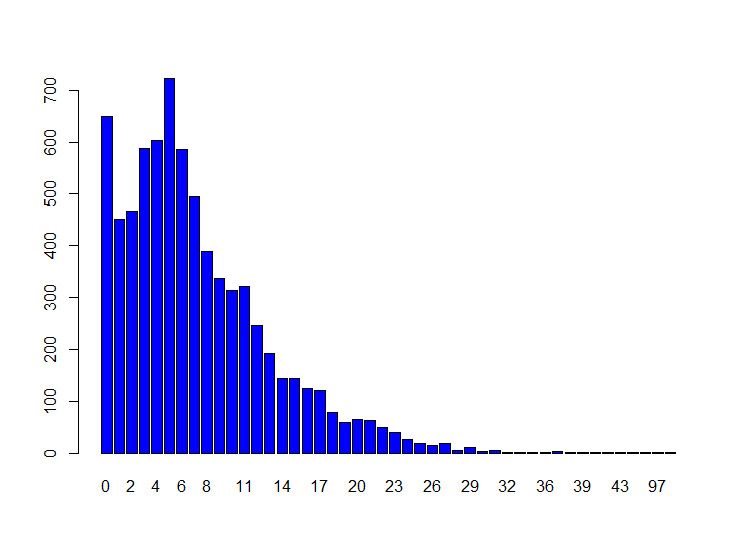
ערכים חסרים: אין

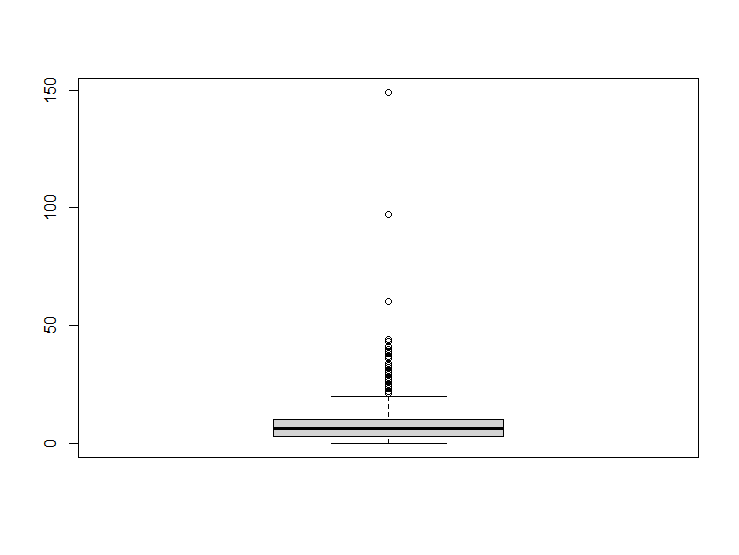
מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 149 | 7.265 | 6 | 6.252 | 3 | 10 |

כמות ערכים חריגים: 278

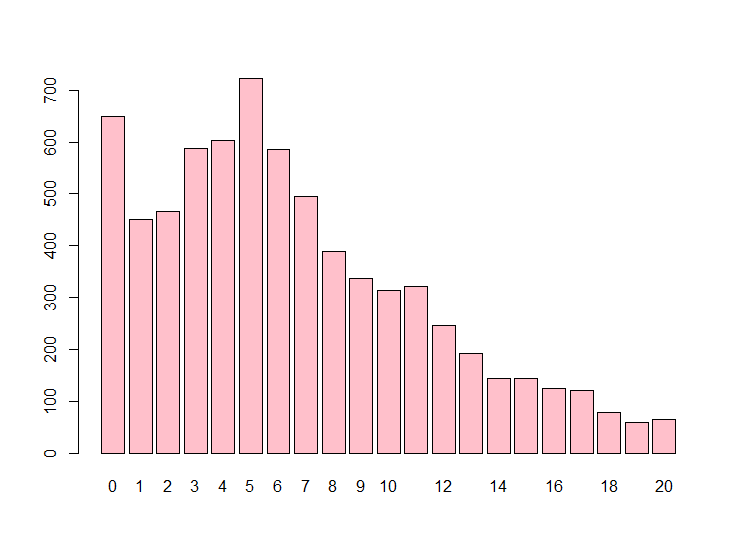
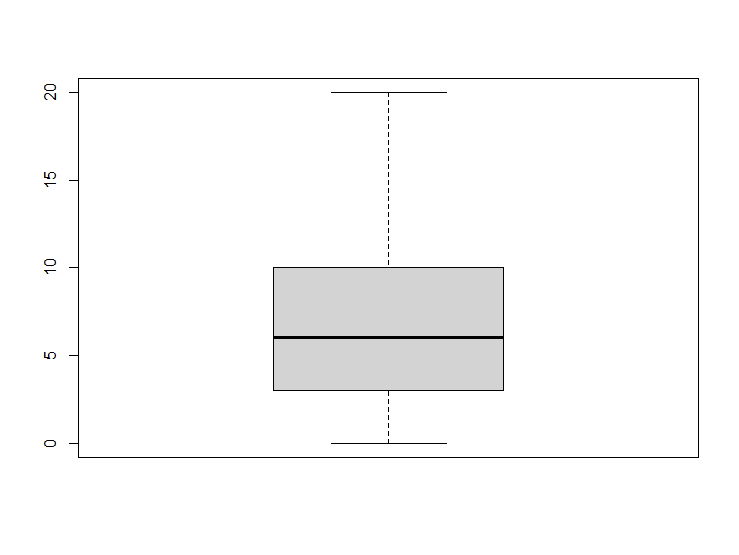
תצוגה גרפית:



ניתוח לאחר הוצאת ה- outliers:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 20 | 6.557 | 6 |  | 3 | 10 |

release\_year

מסקנה – לא להסיר חריגים

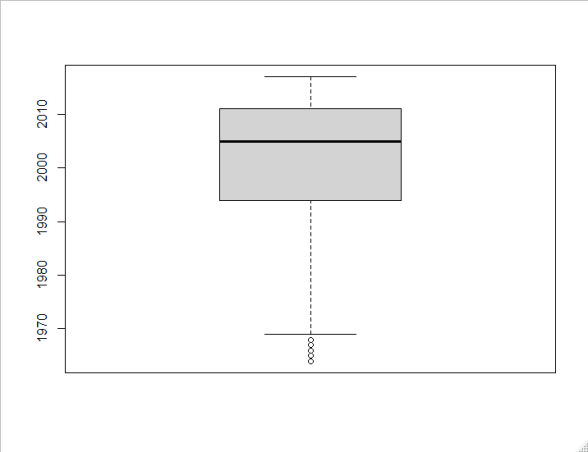
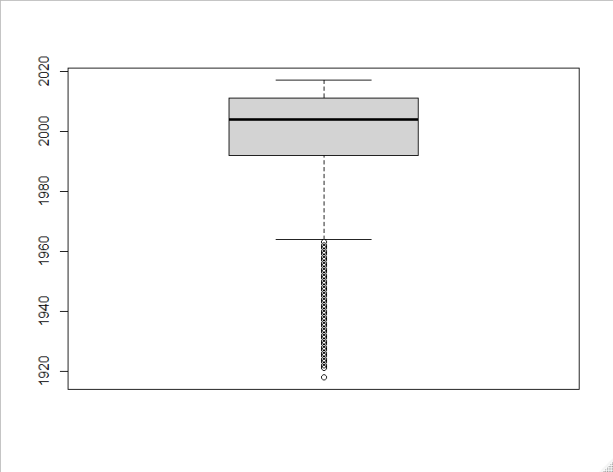
ערכים קיימים: 47,37

ערכים חסרים: 1

מדדי פיזור:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | | **IQR - 25%** | **IQR - 75%** | |
| 1992 | 2017 | 2000 | 2004 | 15.37 | | 1992 | 2011 | |
|  |  |  |  | outliers: 269+1=270 -> NA | | | | |
| 1964 | 2017 | 2000 | 2004 |  | 1994 | | | 2011 |

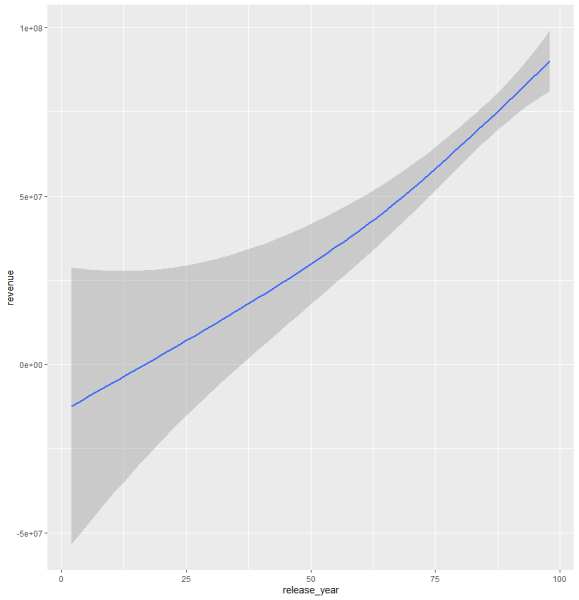
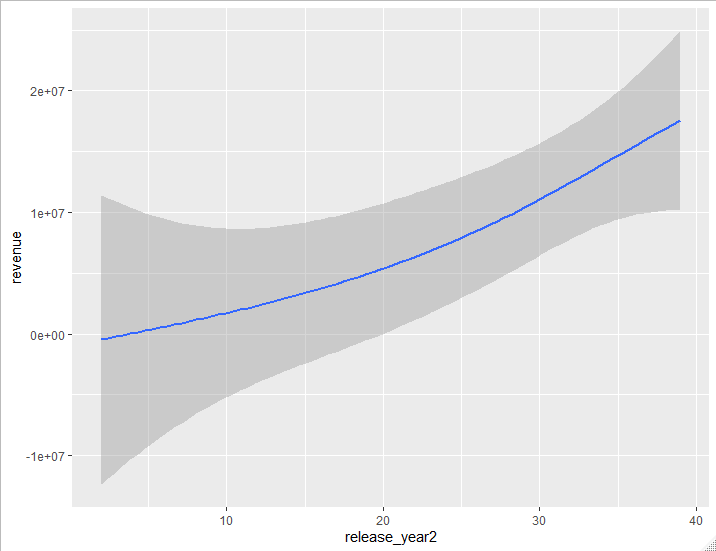
Row data without outliers



Pearson’s correlation

|  |  |  |  |
| --- | --- | --- | --- |
| t = 7.6311 | cor  0.1381951 | p-value = 3.108e-14 | Row data |
| t = 3.1061 | cor  0.3192753 | p-value = 0.002576 | Without outliers |

Row data Correlation- Correlation -without outliers

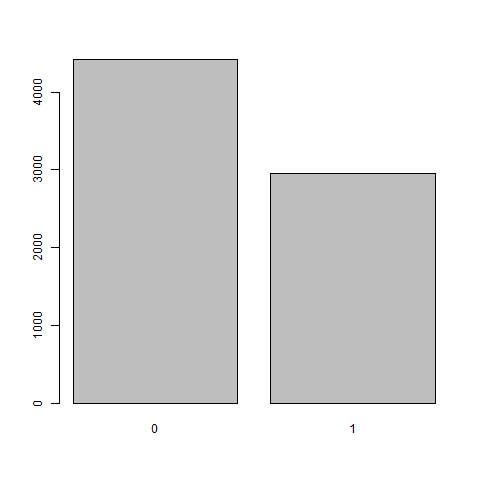


1. release\_month
2. high\_release\_month

ערכים קיימים: 7,375

ערכים חסרים: אין

מספר קטגוריות: 2 (4,427 – לא, 2,948 – כן)

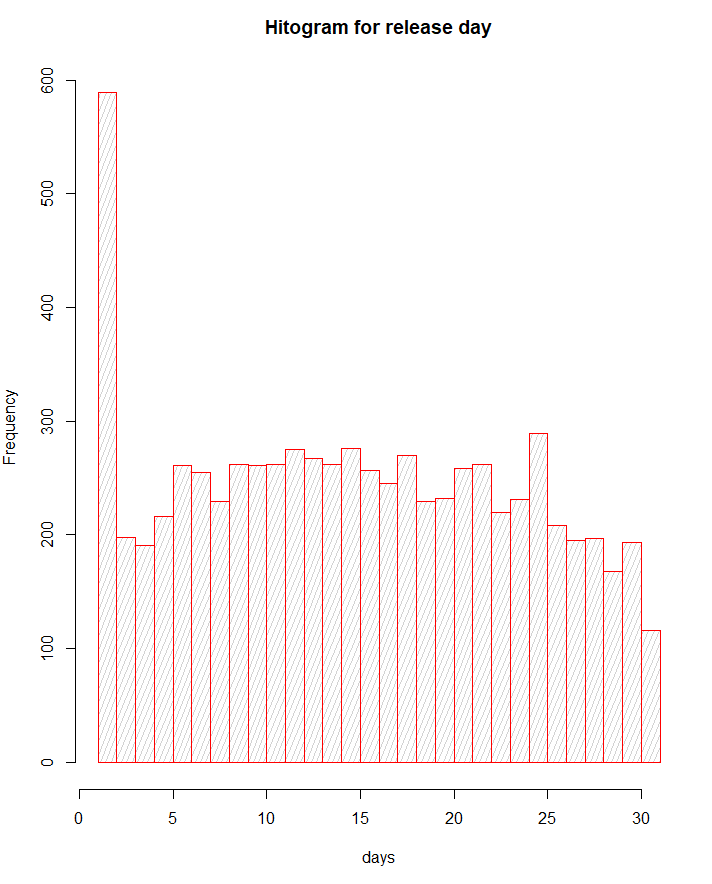


1. release\_day

ערכים קיימים 3,734

ערכים חסרים: 1

מספר קטגוריות: 31



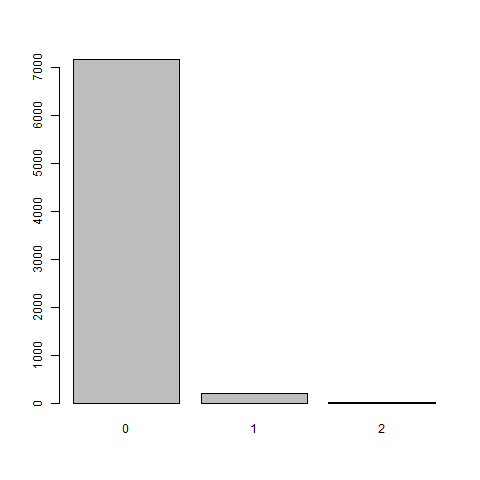
1. seasonality
2. sw\_collection:

ערכים קיימים: 7,375

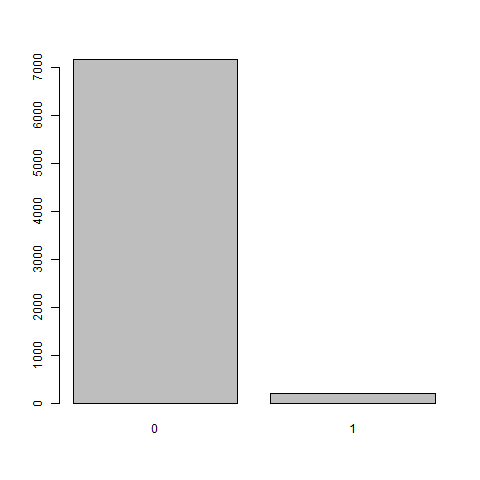
ערכים חסרים: אין

מספר קטגוריות: 3 (7,169 – 0, 195 – 1, 11 -2)

יש להפוך לשתי קטגוריות: 7,169 – לא, 206 – כן



מתוקן עם שתי קטגוריות:



1. producers\_cnt - מסקנה – לא להסיר?

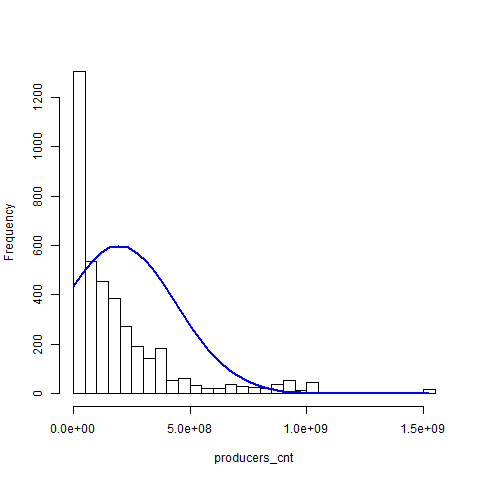
ערכים קיימים: 3,952

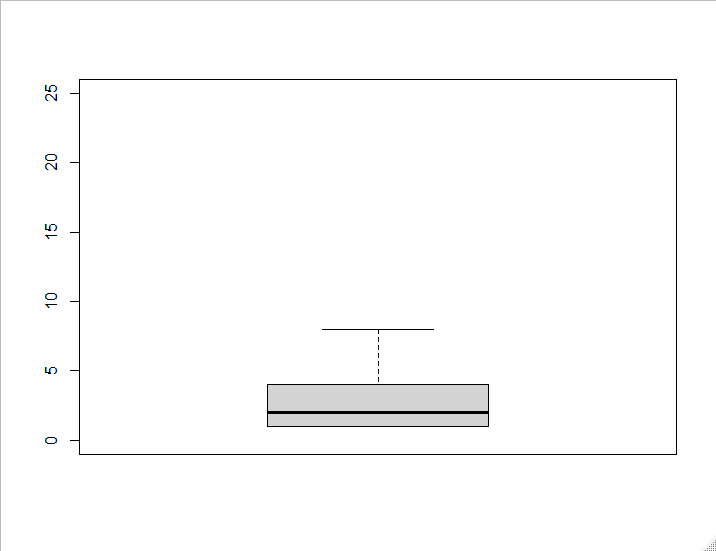
ערכים חסרים: 3,423

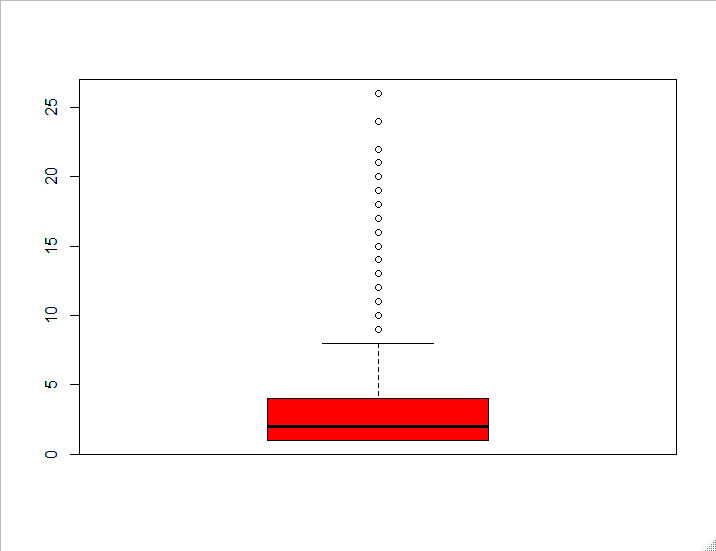
מדדי פיזור:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** | **NA** | |
| **1** | **26** | **2.908** | **2** | **2.143** | **1** | **4** | | **395** |
|  |  |  |  |  |  | **Without outliers** | | |
| **1** | **8** | **2.709** | **2** | **1.658** | **1** | **4** | | **558** |

מספר ה- outliers: 163

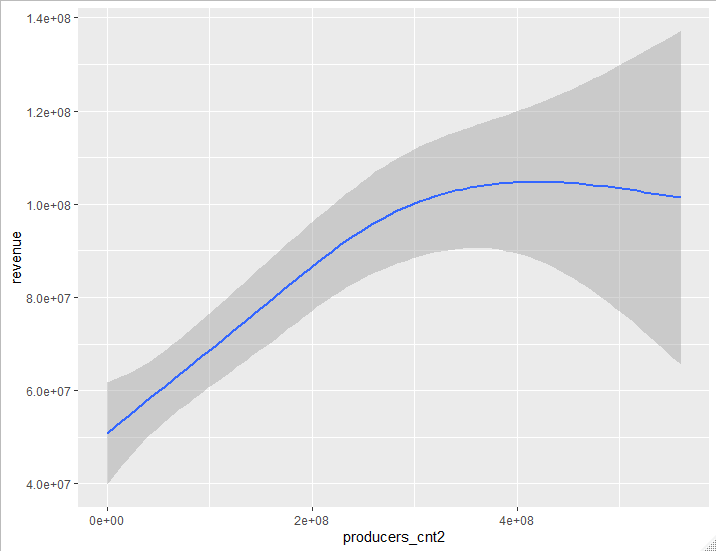
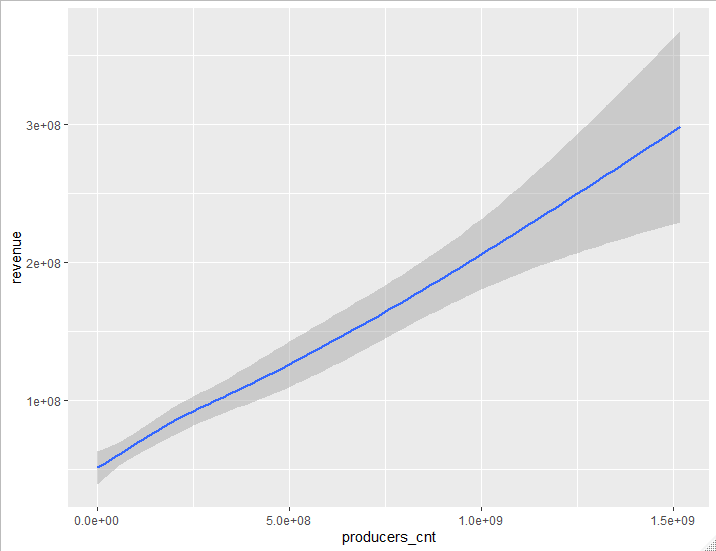
תצוגה גרפית (התפלגות פואסון):

Row data without outliers

Pearson’s correlation

|  |  |  |  |
| --- | --- | --- | --- |
| t = 7.6311 | cor  0.1381951 | p-value = 3.108e-14 | Row data |
| t = 3.1061 | cor  0.3192753 | p-value = 0.002576 | Without outliers |

Row data Correlation- Correlation -without outliers



1. countries\_cnt

|  |  |  |
| --- | --- | --- |
| 1. lang\_US: | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2 (933 - לא, 6,442 – כן) |  |
| 1. lang\_FR | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2 (711 - לא, 6,664) |  |
| 1. lang\_RU |  |  |
| 1. lang\_ES: | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2 (6,816 - לא, 559 – כן) |  |
| 1. lang\_JA | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2 (7171- לא, 204 – כן) |  |

1. actors\_cnt:

ערכים קיימים: 6,688

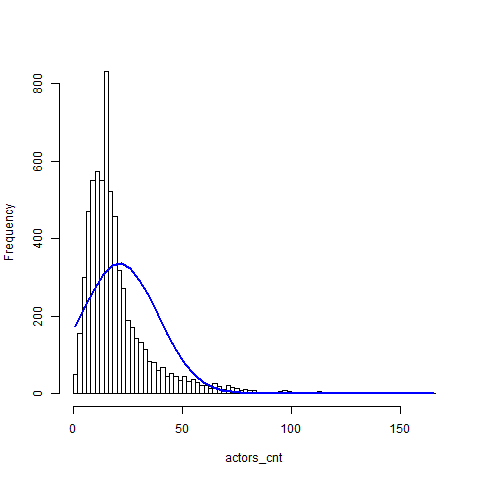
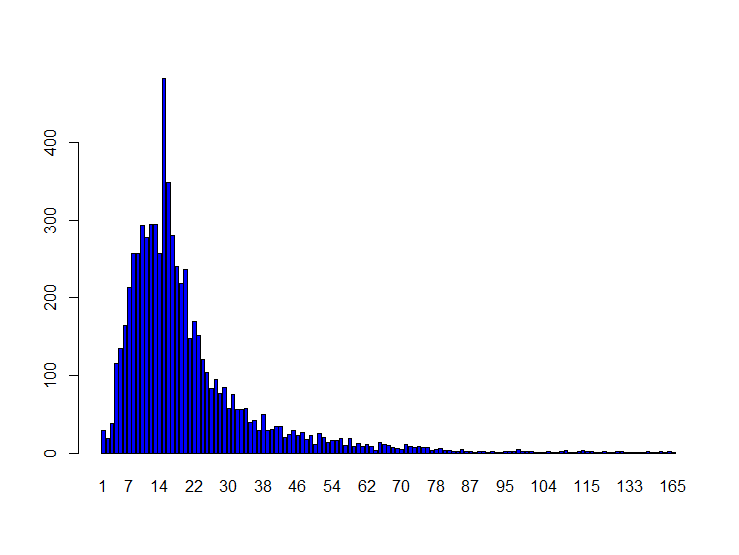
ערכים חסרים: 687

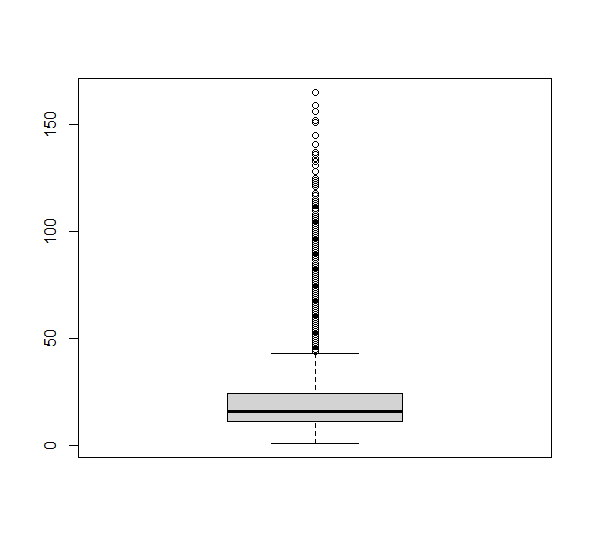
מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 1 | 165 | 21.14 | 16 | 17.53 | 11 | 24 |

מספר ה- outliers: 583

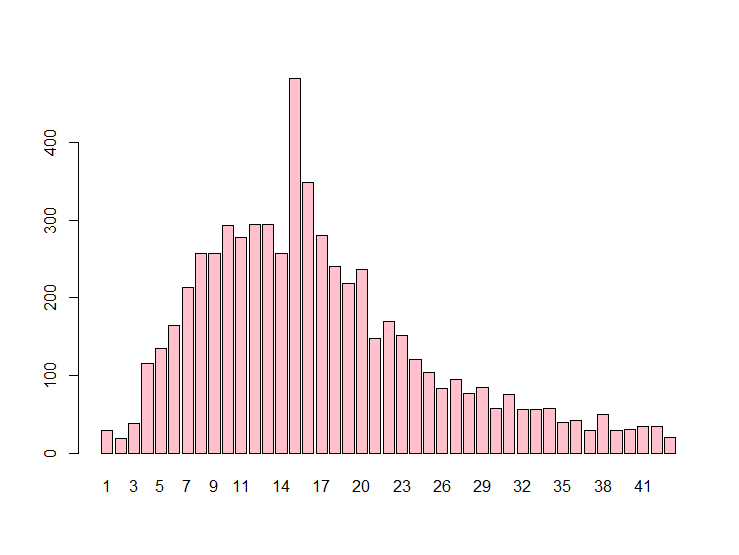
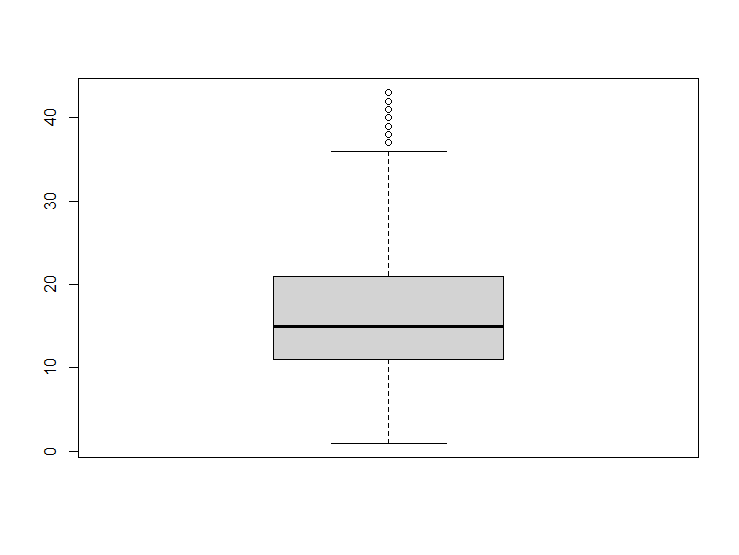
תצוגה גרפית:



ניתוח לאחר הוצאת ה- outliers:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 1 | 43 | 16.85 | 15 |  | 11 | 25 |

1. producers\_cnt:

ערכים קיימים: 6,980

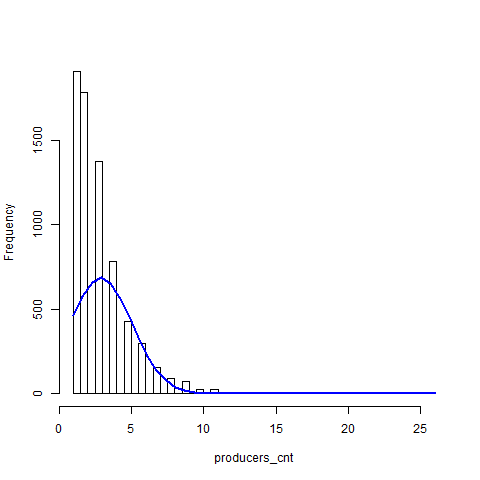
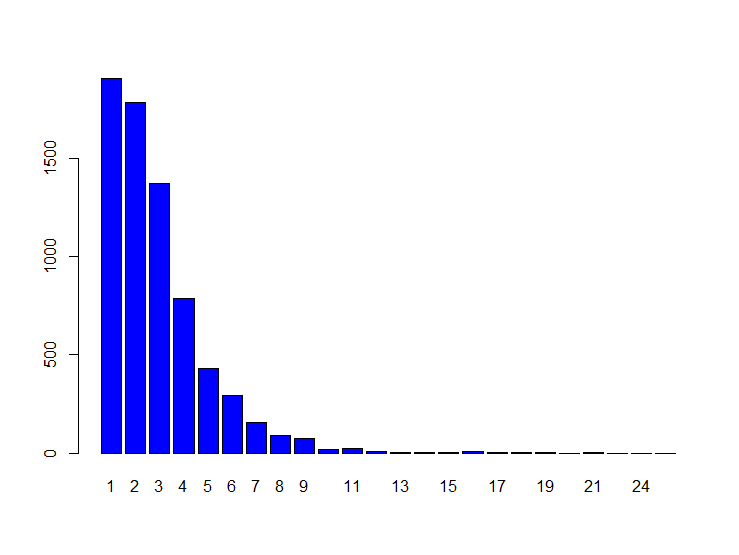
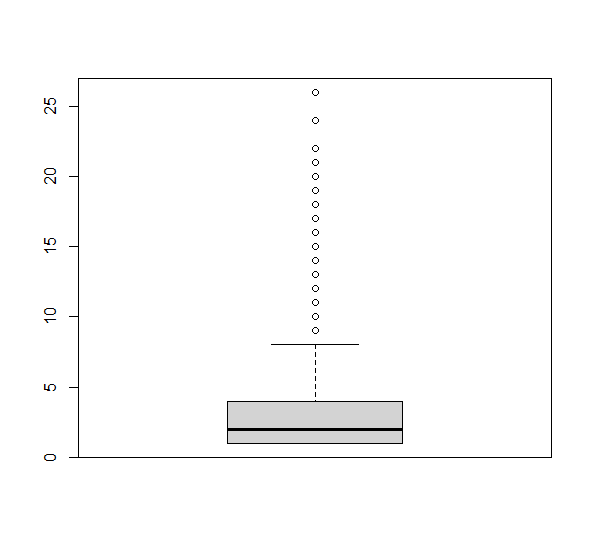
ערכים חסרים: 395

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 1 | 26 | 2.908 | 2 | 2.143 | 1 | 4 |

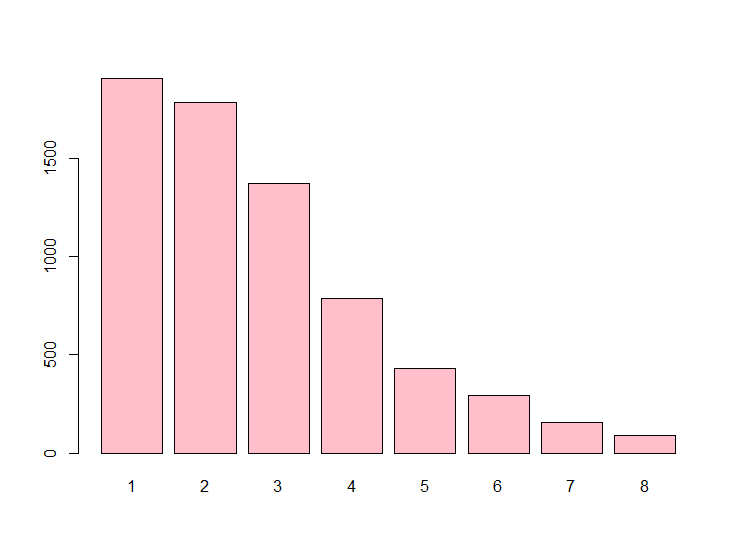
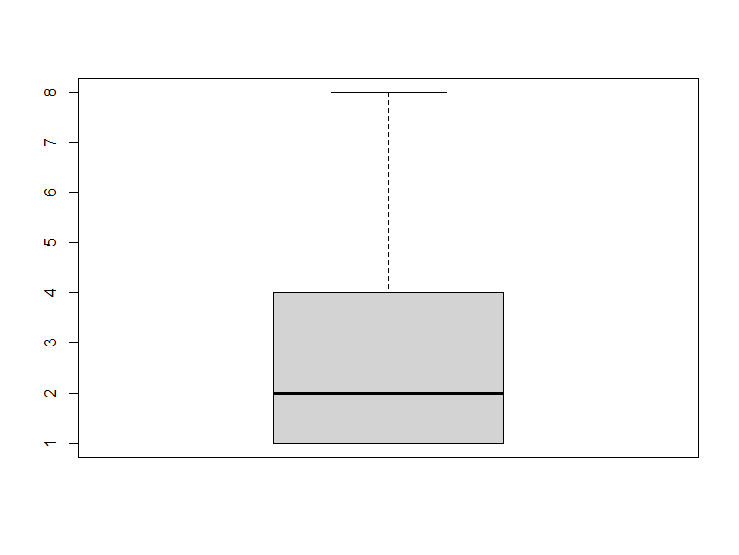
מספר ה- outliers: 163

תצוגה גרפית:

ניתוח לאחר הוצאת ה- outliers:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 1 | 8 | 2.709 | 2 |  | 1 | 4 |

1. actor0\_movies\_cnt
2. actor0\_movies\_5y\_cnt:

ערכים קיימים: 7,375

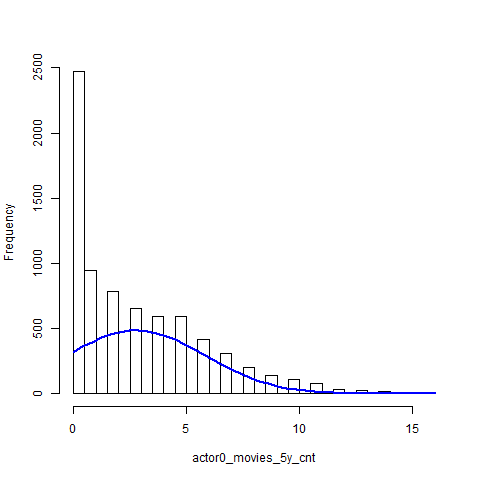
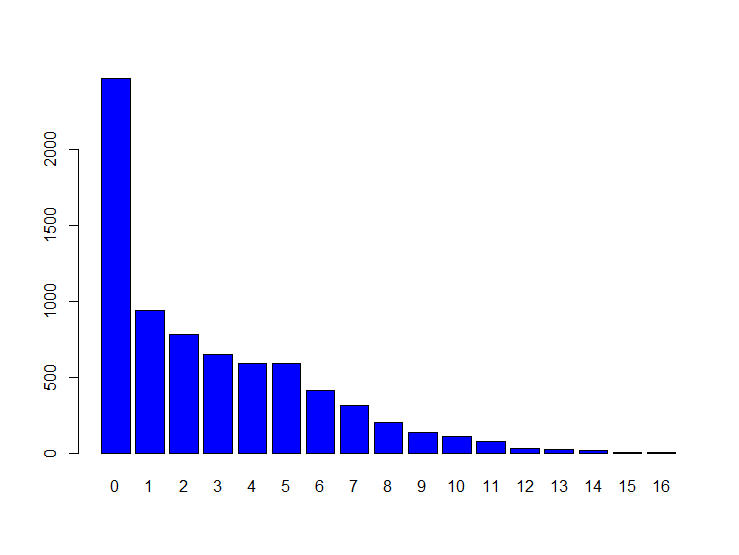
ערכים חסרים: אין

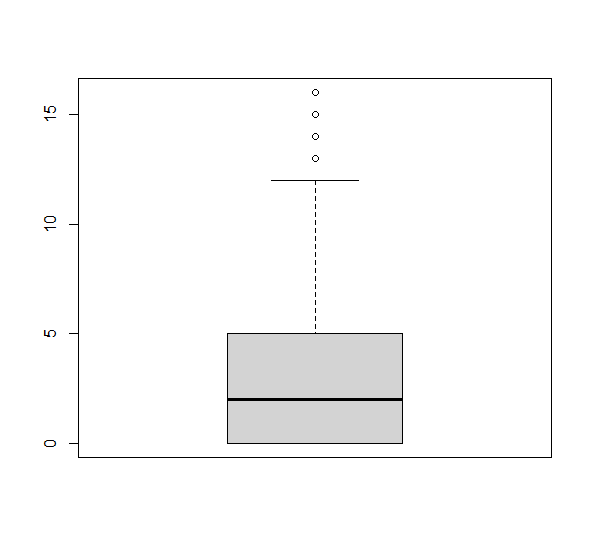
מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 16 | 2.769 | 2 | 3.022 | 0 | 5 |

כמות ערכים חריגים: 163

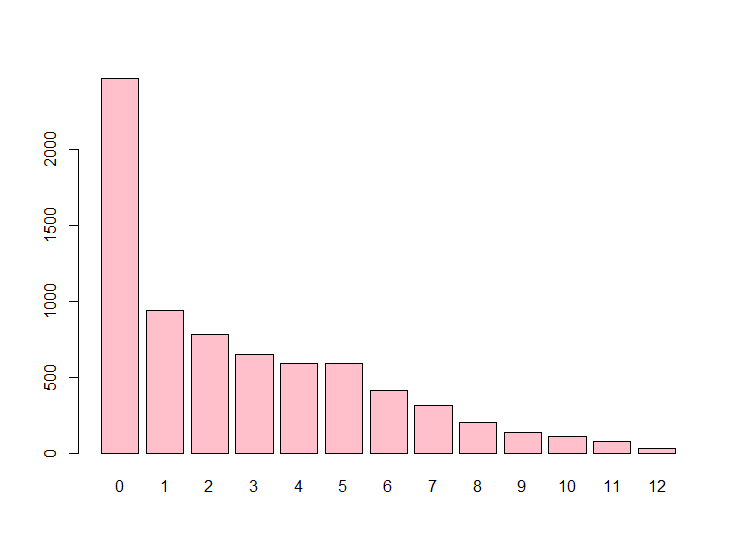
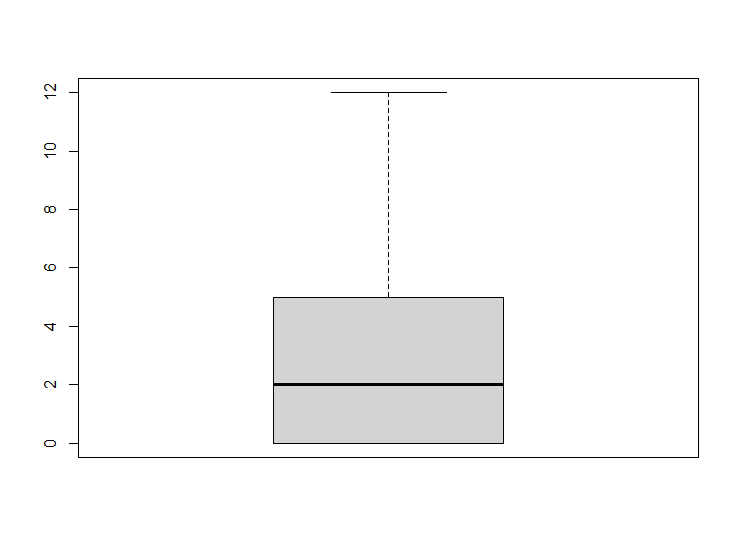
תצוגה גרפית:



ניתוח לאחר הוצאת ה- outliers:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 12 | 2.691 | 2 |  | 0 | 5 |

1. actor1\_movies\_cnt

נשארת קורולציה וההתפלגות משתנה – מסירים outliers

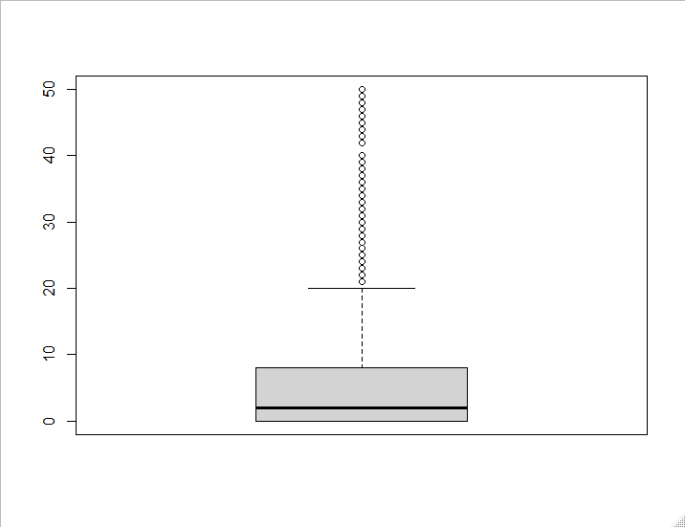
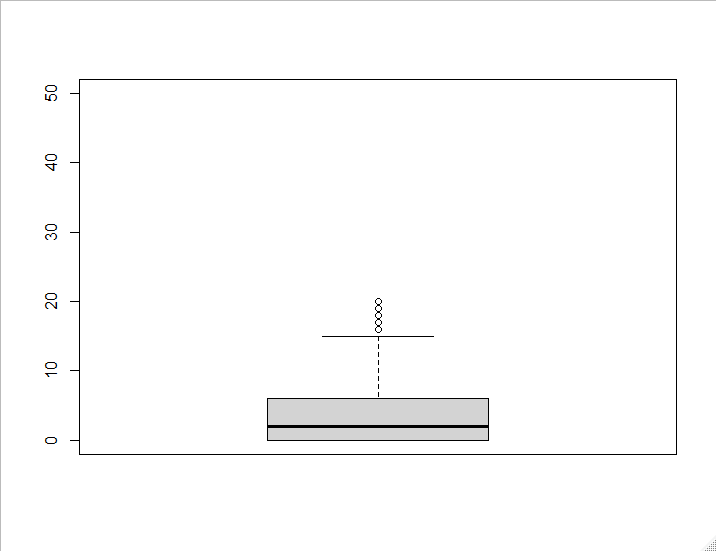
ערכים קיימים: 7,375

ערכים חסרים: אין

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| min | max | mean | median | StdDev | IQR - 25% | IQR - 75% |
| 0 | **67** | **5.405** | **2** | **7.593** | **0** | **8** |
|  |  |  |  | **Without outliers** | | |
| 0 | **20** | **4.066** | **2** | **5.054** | **0** | **6** |

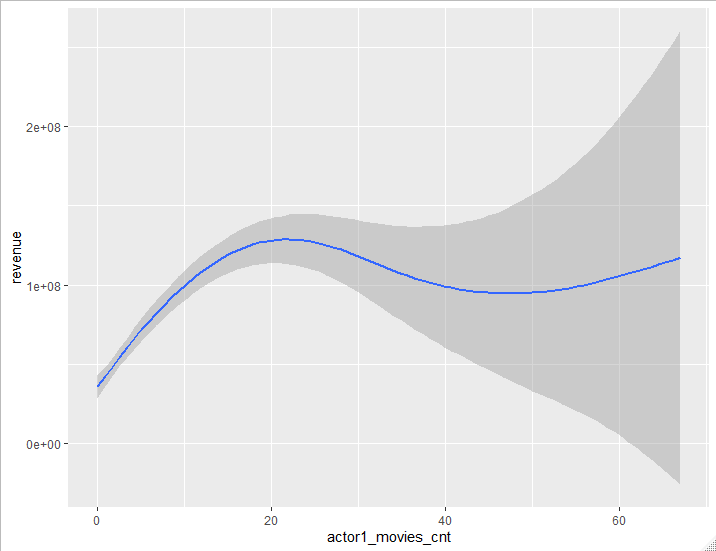
מספר ה- outliers: 411

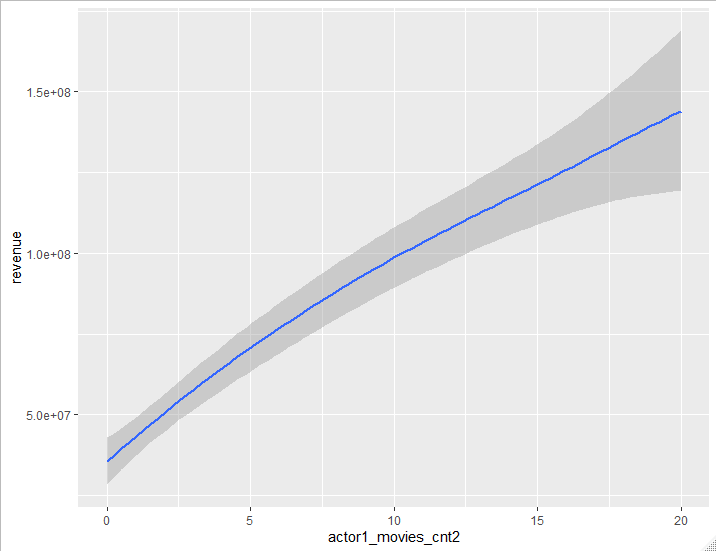
Row data without outliers

spearman’s correlation

|  |  |  |  |
| --- | --- | --- | --- |
| S = 3033084833 | rho 0.3212408 | p-value < 2.2e-16 | Row data |
| S = 2512121969 | rho 0.3163119 | p-value < 2.2e-16 | Without outliers |

Row data Correlation- Correlation -without outliers





1. actor1\_movies\_5y\_cnt
2. actor2\_movies\_cnt:

ערכים קיימים: 7,375

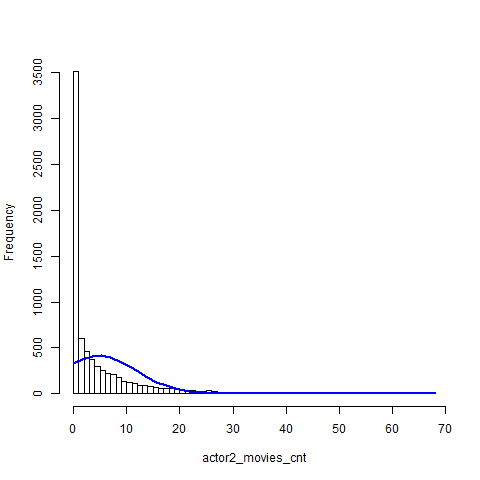
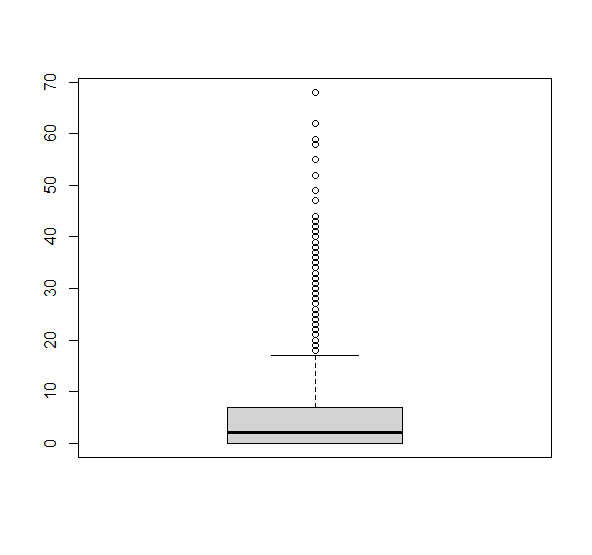
ערכים חסרים: אין

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 68 | 4.773 | 2 | 7.102 | 1 | 7 |

מספר ה- outliers: 35

תצוגה גרפית:

1. actor2\_movies\_5y\_cnt

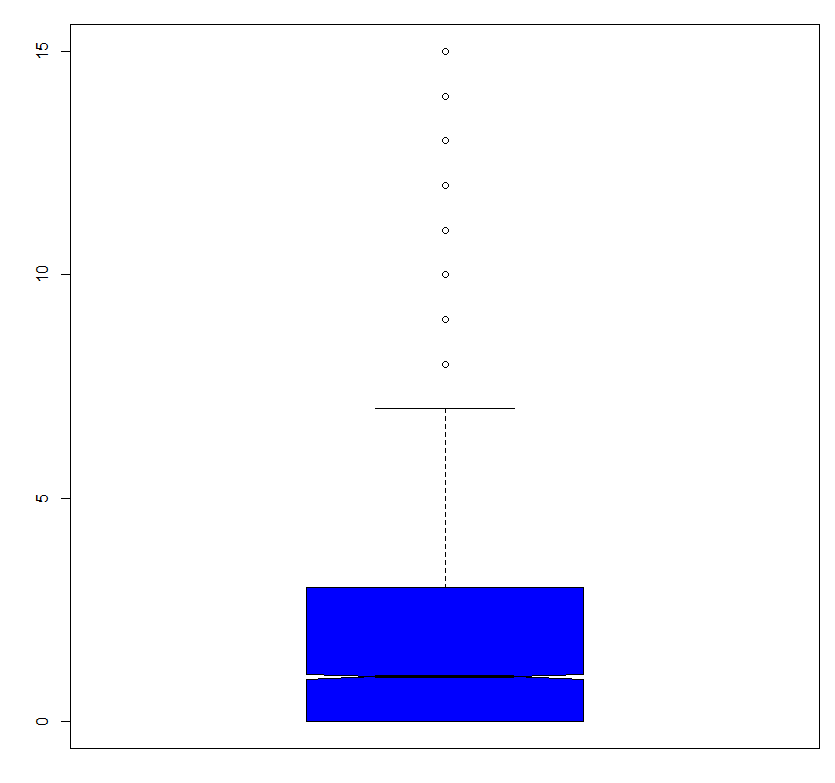
ערכים קיימים: 7,375

ערכים חסרים: אין

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| **0** | **15** | **1.915** | **1** | **2.559** | **0** | **3** |

מספר ה- outliers: 345



|  |  |  |
| --- | --- | --- |
| 1. sw\_female\_actor0 |  |  |
| 1. sw\_female\_actor1: | ערכים קיימים: 6,593  ערכים חסרים: 782  מספר קטגוריות: 2 (2,500 - כן, 4,093 – לא)  הערות: ערכים חסרים – אין שחקן מספר 1 בסרט. מה נעשה עם החוסר? |  |
| 1. sw\_female\_actor2 | ערכים קיימים: 6587  ערכים חסרים: 788  מספר קטגוריות: 2 (2226 - כן, 4361 – לא) |  |
| 1. sw\_male\_actor0 |  |  |
| 1. sw\_male\_actor1: | ערכים קיימים: 6,593  ערכים חסרים: 782  מספר קטגוריות: 2 (3,409 - כן, 3,184 – לא) |  |
| 1. sw\_male\_actor2 | ערכים קיימים: 7,372  ערכים חסרים: 788  מספר קטגוריות: 2 (3581 - 1, 3006 – 0) |  |
| 1. actor0\_prev\_revenue |  |  |

1. actor1\_prev\_revenue:

ערכים קיימים: 4,174

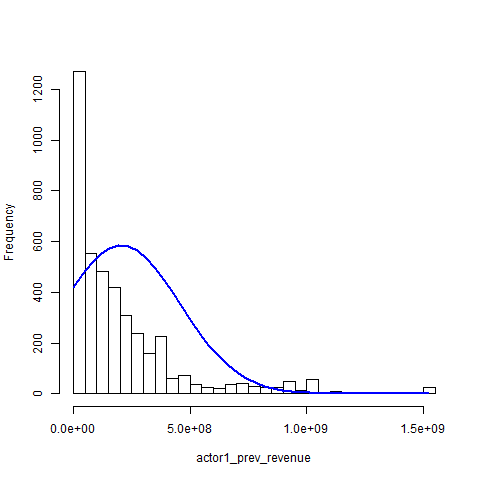
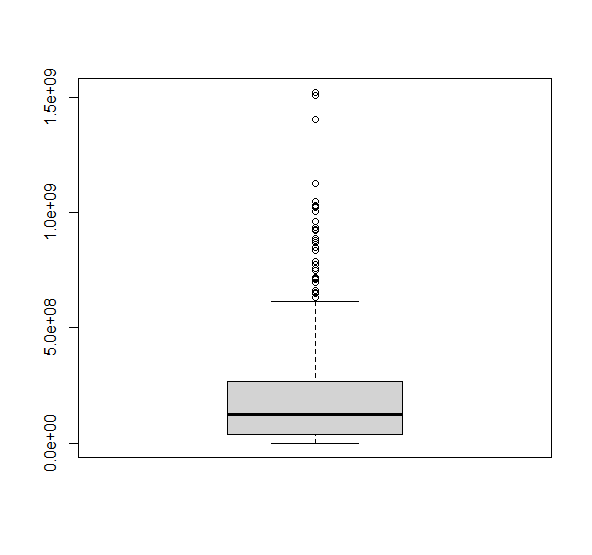
ערכים חסרים: 3,201

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 68 | 4.773 | 2 | 7.102 | 1 | 7 |

מספר ה- outliers: ?

תצוגה גרפית:

1. actor2\_prev\_revenue

ערכים קיימים: 3952

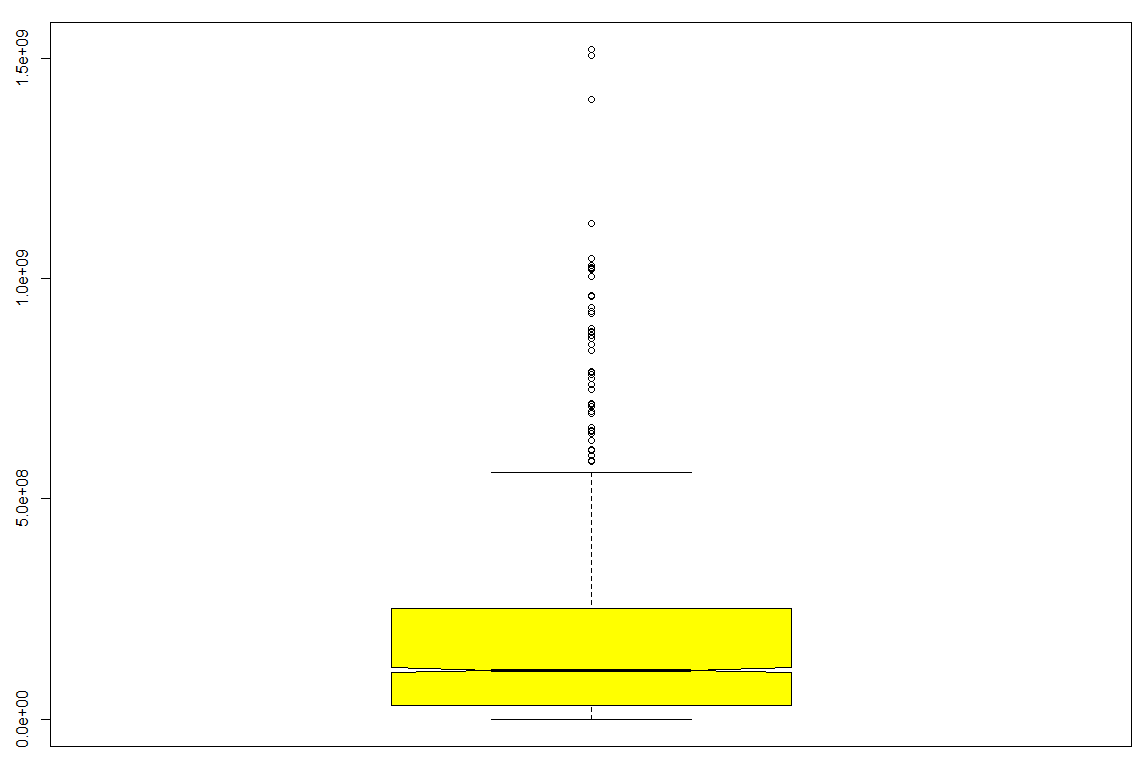
ערכים חסרים: 3423

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| **1** | **1.52E+09** | **1.96E+08** | **1.11E+08** | **2.46E+08** | **3.09E+07** | **2.51E+08** |

מספר ה- outliers: 331

תצוגה גרפית:



1. director\_movies\_cnt:
2. director\_movies\_5y\_cnt:

ערכים קיימים: 7,375

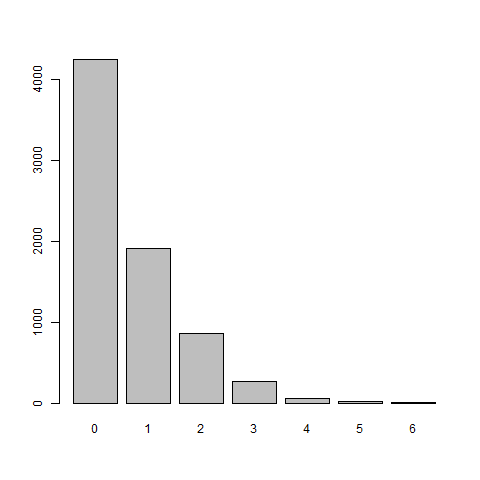
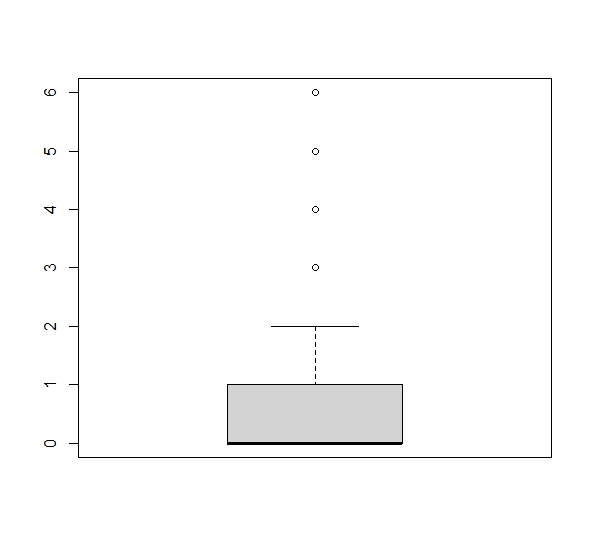
ערכים חסרים: אין

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 68 | 4.773 | 2 | 7.102 | 1 | 7 |

מספר ה- outliers: 4

תצוגה גרפית:

|  |  |  |
| --- | --- | --- |
| genre |  | תצוגה גרפית |
| 1. \_adventure | ערכים קיימים: 7,375  ערכים חסרים: אין  קטגוריה כן (1) -  קטגוריה לא (0) |  |
| 1. \_fantasy |  |  |
| 1. animation: | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2  (382 - כן,  6,993 – לא) |  |
| 1. \_drama | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2 (3676 - כן,  3699 – לא) |  |
| 1. \_horror |  |  |
| 1. action | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2 (1,735 - כן, 5,640 – לא) |  |
| 1. comedy | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2 (2605 - כן,  4770 – לא) |  |
| 1. history |  |  |
| 1. western | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2 (117 - כן, 7,258 – לא) |  |
| 1. thriller | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2  (1869- כן,  5506 – לא) |  |
| 1. crime |  |  |
| 1. documentary | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2 (221 - כן, 7,154 – לא) |  |
| 1. science fiction | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2  (744 - כן,  6631 – לא) |  |
| 1. mystery |  |  |
| 1. music | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2  (267 - כן,  7,108 – לא) |  |
| 1. romance | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2 (1435 - כן,  5940 – לא) |  |
| 1. family |  |  |
| 1. war | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2  (243 - כן,  7,132 – לא) |  |
| 1. foriegn | ערכים קיימים: 7,375  ערכים חסרים: אין  מספר קטגוריות: 2  (84 - כן,  7291– לא) |  |

depart\_Art

1. depart\_Camera:

ערכים קיימים: 7,341

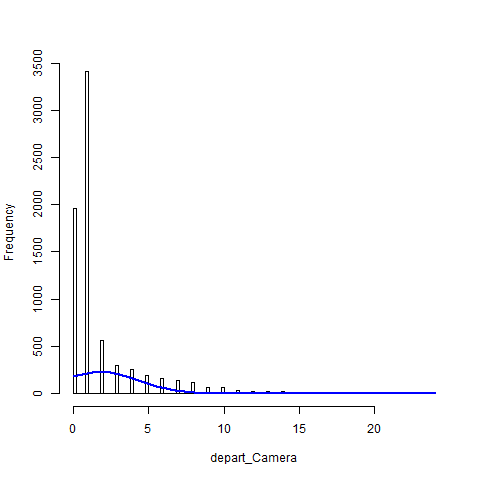
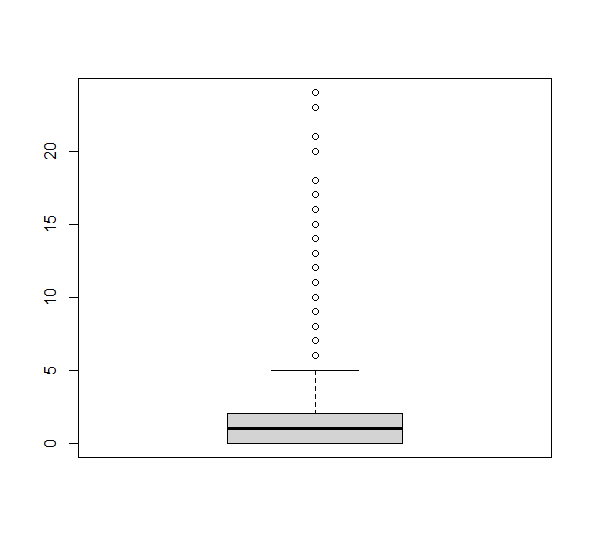
ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 24 | 1.802 | 1 | 6.892 | 0 | 2 |

מספר ה- outliers: ??

תצוגה גרפית:

1. deaprt\_crew

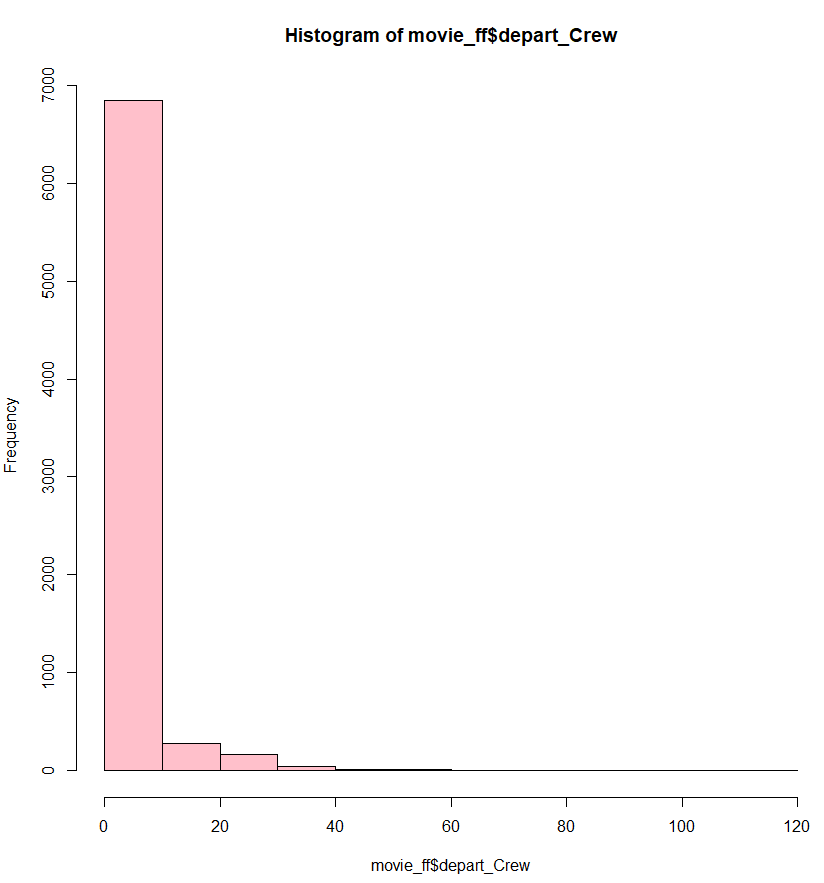
ערכים קיימים: 7,341

ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 119 | 2.388 | 0 | 5.875 | 0 | 2 |

מספר ה- outliers: 906



1. depart\_custom\_mkup
2. depart\_Directing:

ערכים קיימים: 7,341

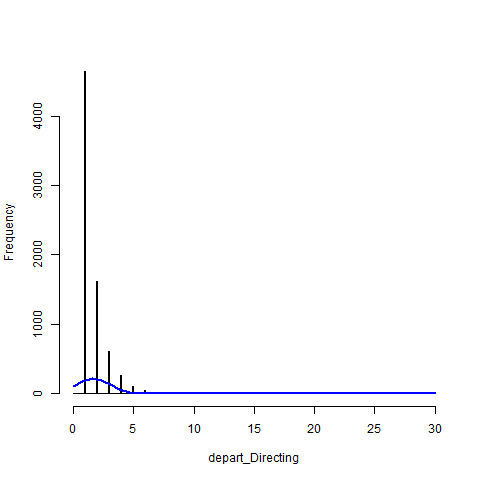
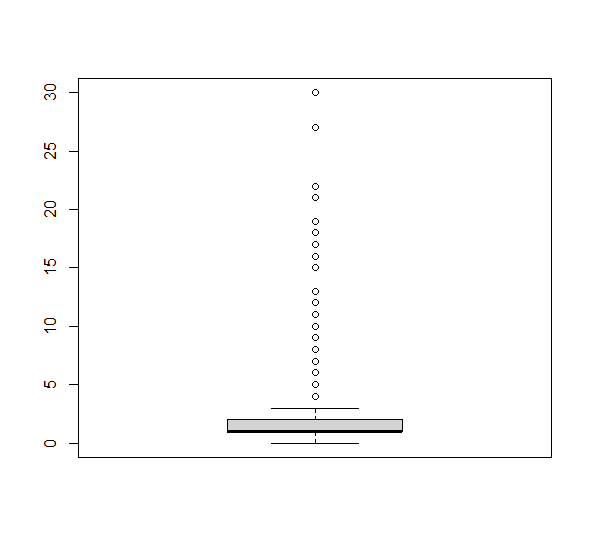
ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 30 | 1.665 | 1 | 6.109 | 1 | 2 |

מספר ה- outliers: ?

תצוגה גרפית:

1. depart\_Editing:

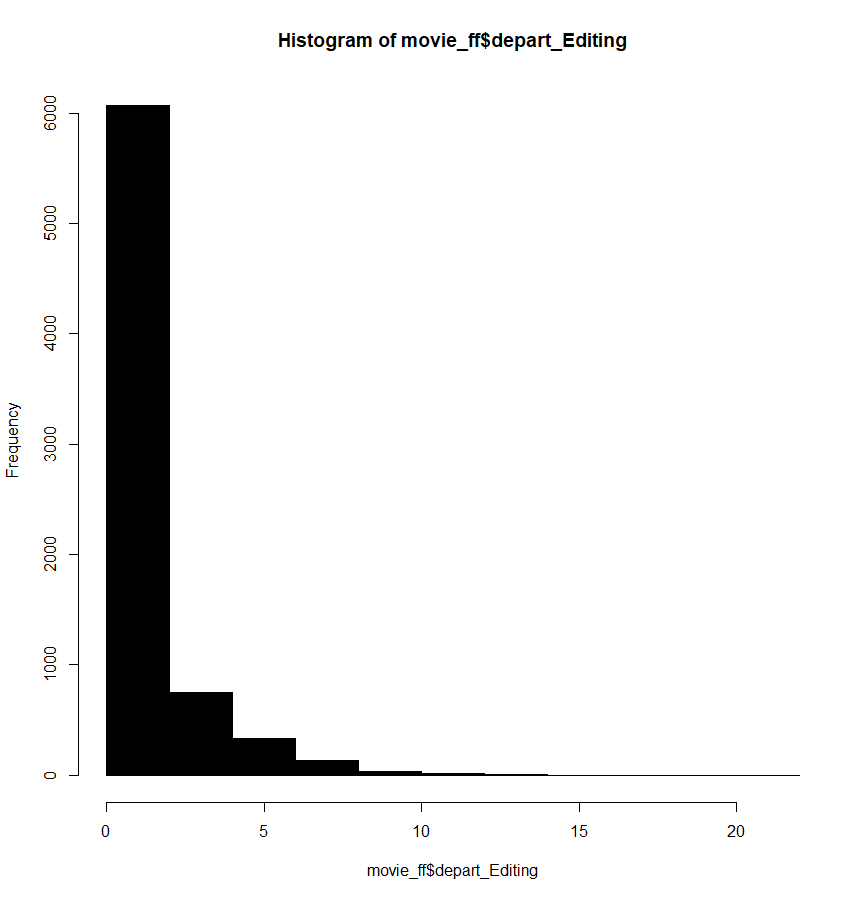
ערכים קיימים: 7,341

ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 21 | 1.516 | 1 | 1.738 | 0 | 2 |

מספר ה- outliers: 327



1. depart\_Lightning:
2. depart\_Production:

ערכים קיימים: 7,341

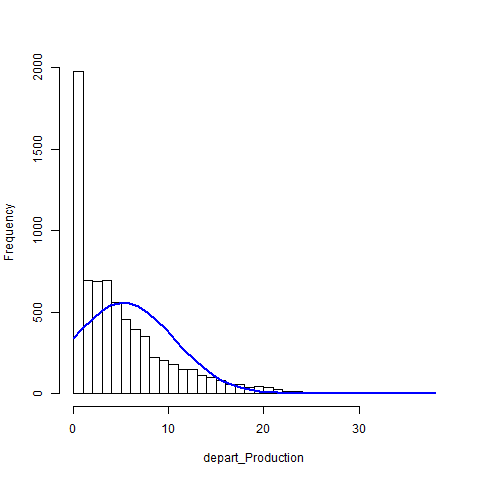
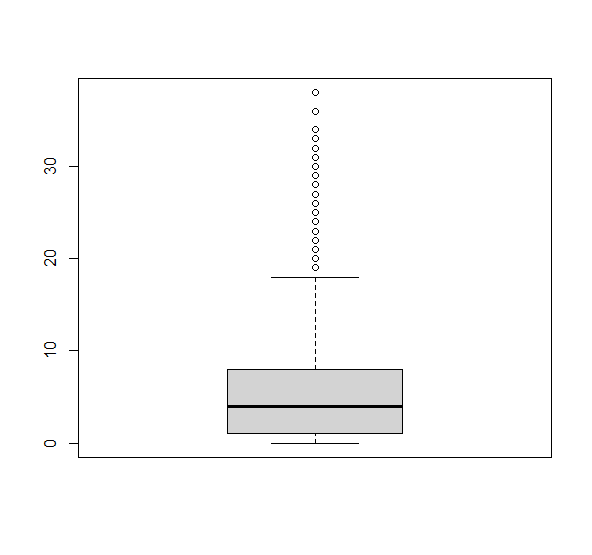
ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 38 | 5.301 | 4 | 14.07 | 1 | 8 |

מספר ה- outliers: ?

תצוגה גרפית:

1. depart\_Sound

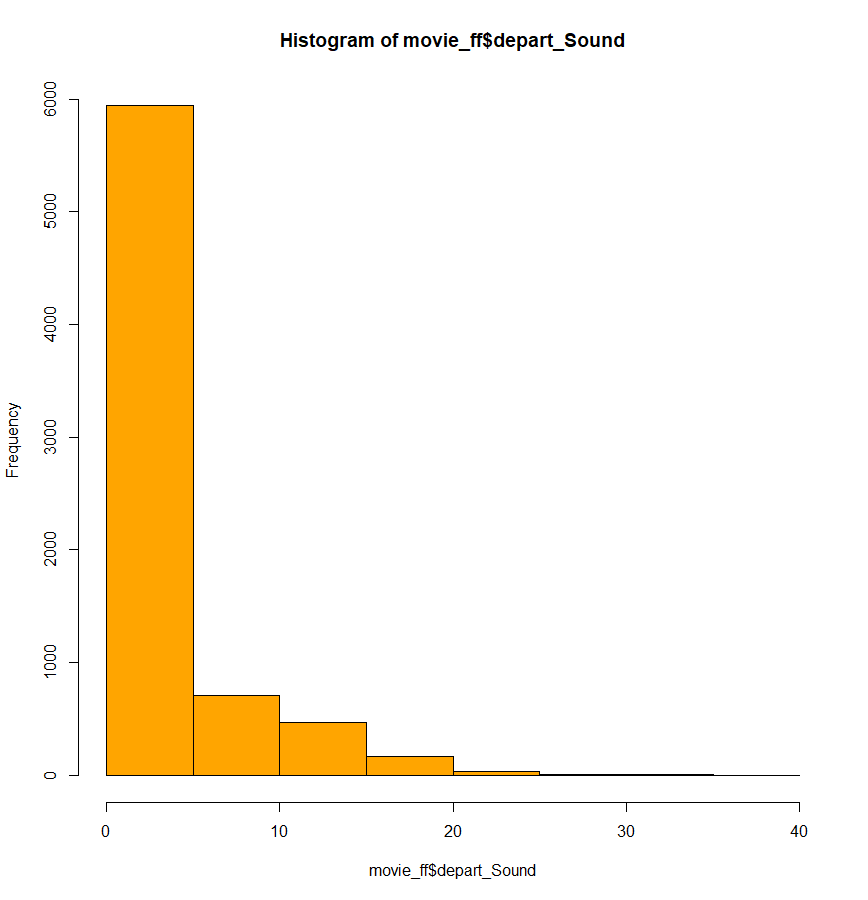
ערכים קיימים: 7,341

ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 40 | 3.064 | 1 | 4.516 | 1 | 3 |

מספר ה- outliers: 2890



1. depart Visual effect
2. depart\_Writing:

ערכים קיימים: 7,341

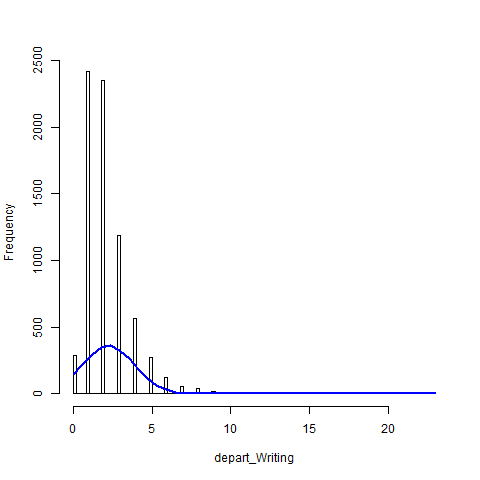
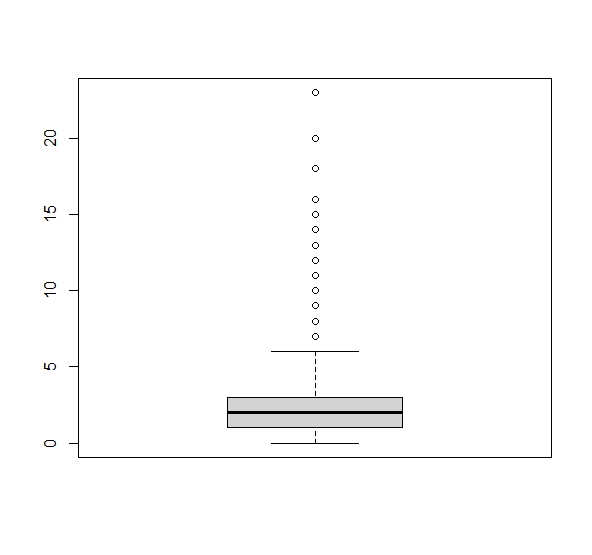
ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 23 | 2.222 | 2 | 5.615 | 1 | 3 |

מספר ה- outliers: 415

תצוגה גרפית:

1. Depart\_art\_female

ערכים קיימים: 7,341

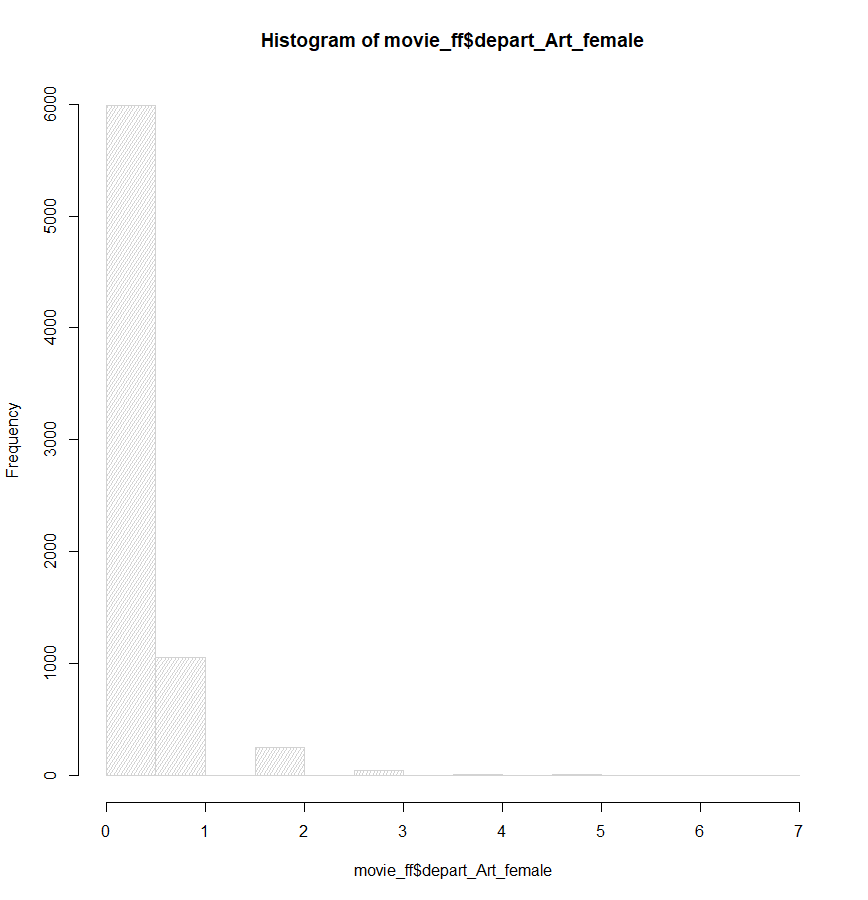
ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 7 | 0.2351 | 0 | 0.558095 | 0 | 0 |

מספר ה- outliers: 1055

תצוגה גרפית:



1. depart\_camera\_female
2. depart\_Crew\_female:

ערכים קיימים: 7,341

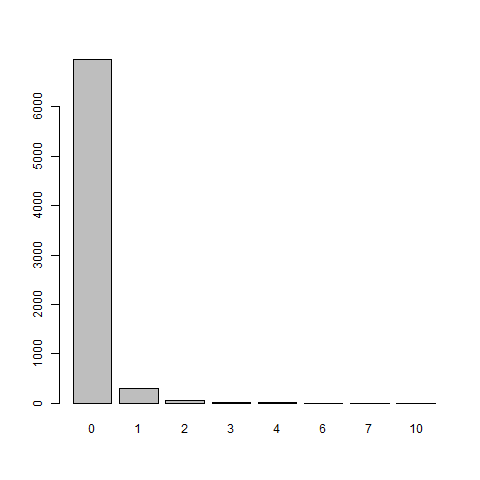
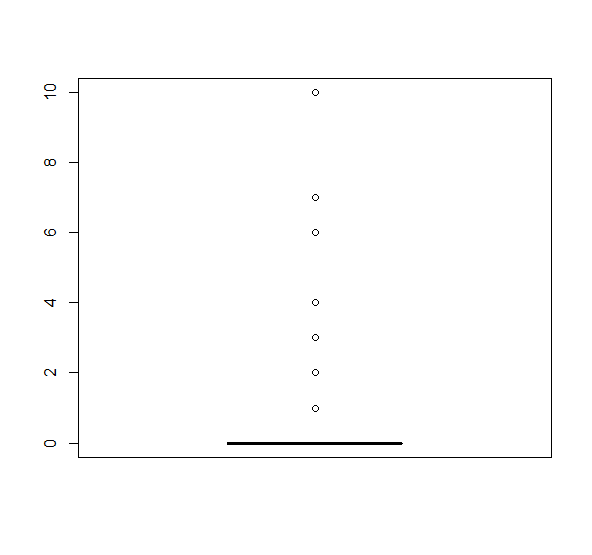
ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 10 | 0.06443 | 0 |  | 0 | 0 |

מספר ה- outliers: ?

תצוגה גרפית:

1. depart\_custom\_mkup\_female

ערכים קיימים: 7,341

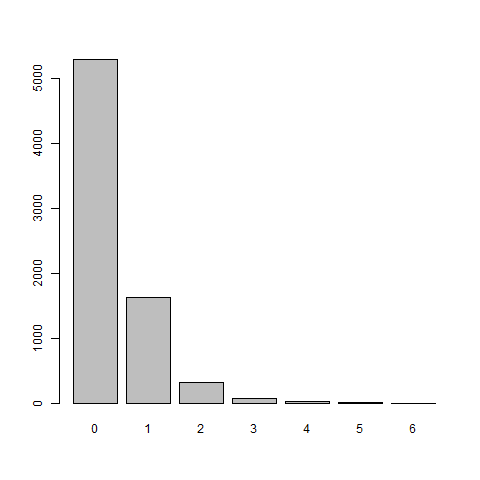
ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 6 | 0.3535 | 0 | 0.651521 | 0 | 1 |

מספר ה- outliers: 1631

תצוגה גרפית:



1. depart\_directing\_femal
2. depart\_Editing\_female:

ערכים קיימים: 7,341

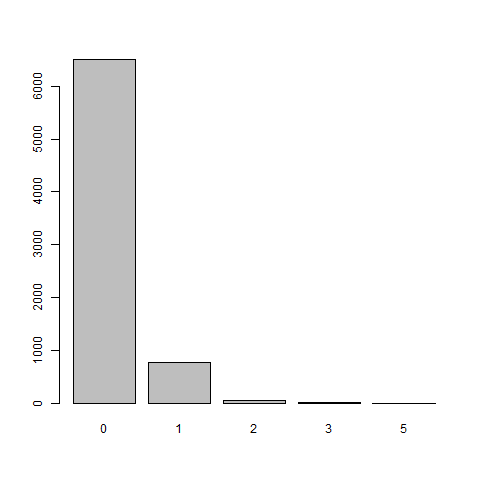
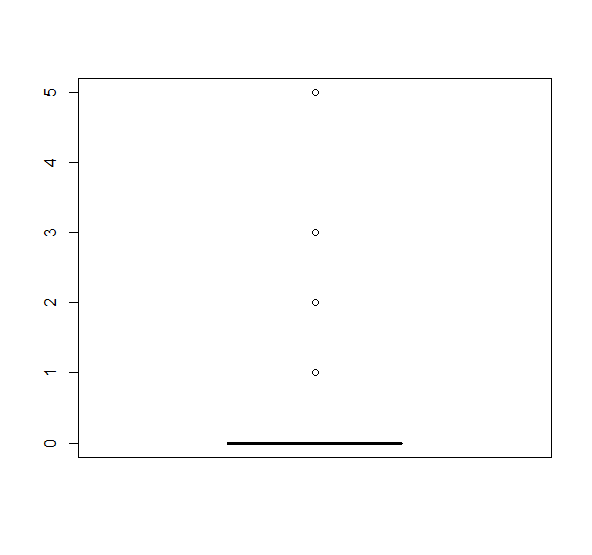
ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 5 | 0.1216 | 0 |  | 0 | 0 |

מספר ה- outliers: ?

תצוגה גרפית:

1. depart\_lighting\_female

ערכים קיימים: 7,341

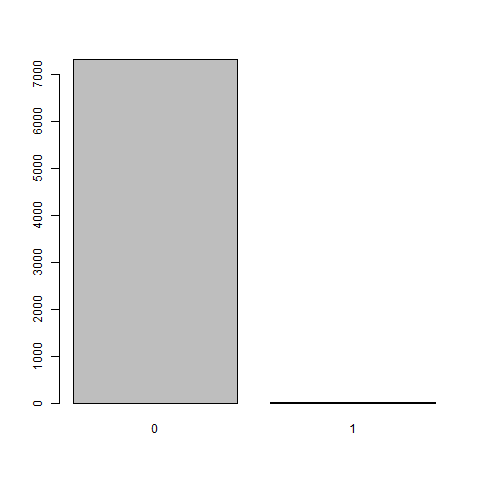
ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 1 | 0.00068 | 0 | 0.026091 | 0 | 0 |

מספר ה- outliers: 5

תצוגה גרפית:



1. depart\_production\_female
2. depart\_Sound\_female:

ערכים קיימים: 7,341

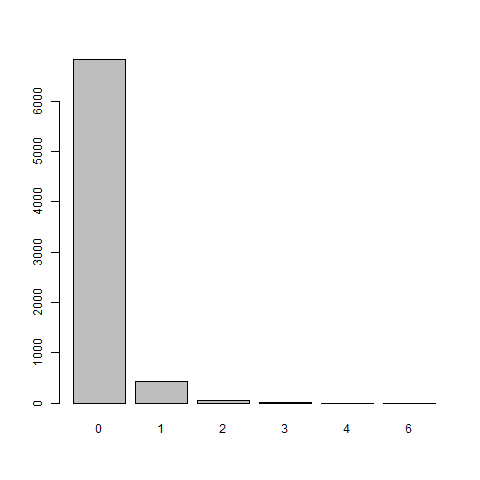
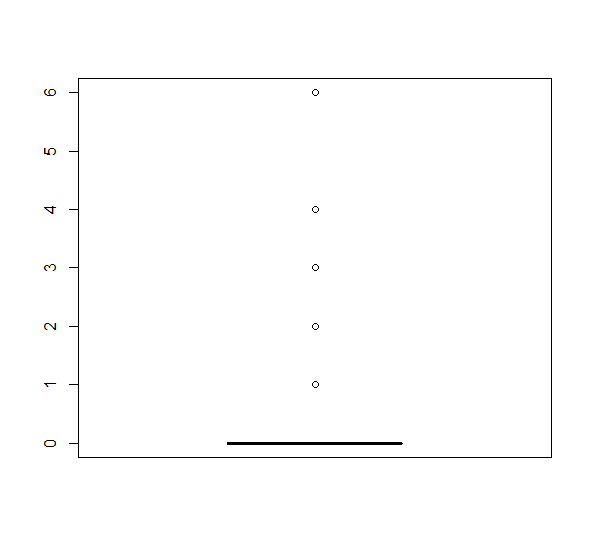
ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 6 | 0.07914 | 0 |  | 0 | 0 |

כמות ערכים חריגים:

תצוגה גרפית:

הערה כללית: האם אפסים הם אמיתיים או חוסרים?

1. depart\_visual\_effects\_female

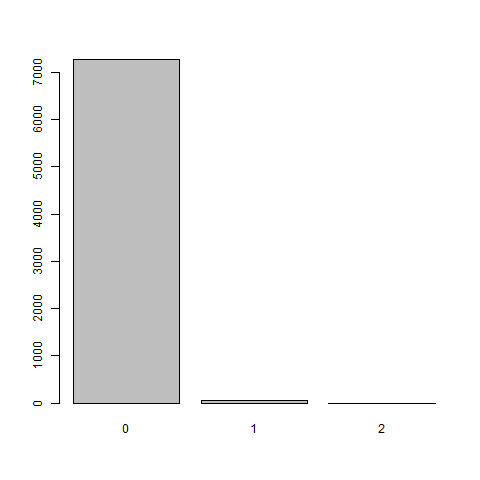
ערכים קיימים: 7,341

ערכים חסרים: 34

מדדי פיזור:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Min** | **Max** | **Mean** | **Median** | **StdDev** | **IQR - 25%** | **IQR - 75%** |
| 0 | 2 | 0.00845 | 0 | 0.094448 | 0 | 0 |

מספר ה- outliers: 58



1. deaprt\_writing\_female
2. job interraction