שיטות מחקר תרגיל 3

# שאלה 1

1. ניתן לראות בקובץ ה-R המצורף.
2. ניתן לראות בקובץ ה-R המצורף.

|  |  |  |
| --- | --- | --- |
| מודל | p-value | גודל אפקט |
| רגרסיה פשוט עם מנבא x | 0.1152 | 0.1044 |
| רגרסיה מרובה עם מנבאים x,x2 | 0.005901 | 0.3729 |
| רגרסיה מרובה עם מנבאים x,x2,x3 | 0.01257 | 0.3967 |
| רגרסיה מרובה עם מנבאים x,x2,x3,x4 | 0.03076 | 0.3988 |
| רגרסיה מרובה עם מנבאים x,x2,x3,x4,x5 | 0.06213 | 0.4026 |

|  |  |
| --- | --- |
| מודל | SSE |
| רגרסיה פשוט עם מנבא x | 14.2 |
| רגרסיה מרובה עם מנבאים x,x2 | 9.94 |
| רגרסיה מרובה עם מנבאים x,x2,x3 | 9.57 |
| רגרסיה מרובה עם מנבאים x,x2,x3,x4 | 9.53 |
| רגרסיה מרובה עם מנבאים x,x2,x3,x4,x5 | 9.47 |

1. ניתן לראות כי ערך ה-SSE הנמוך ביותר הוא של המודל רגרסיה מרובה עם מנבאים x,x2,x3,x4,x5.
2. Chart, line chart

   Description automatically generated

Chart, line chart

Description automatically generated

Chart

Description automatically generated

Chart

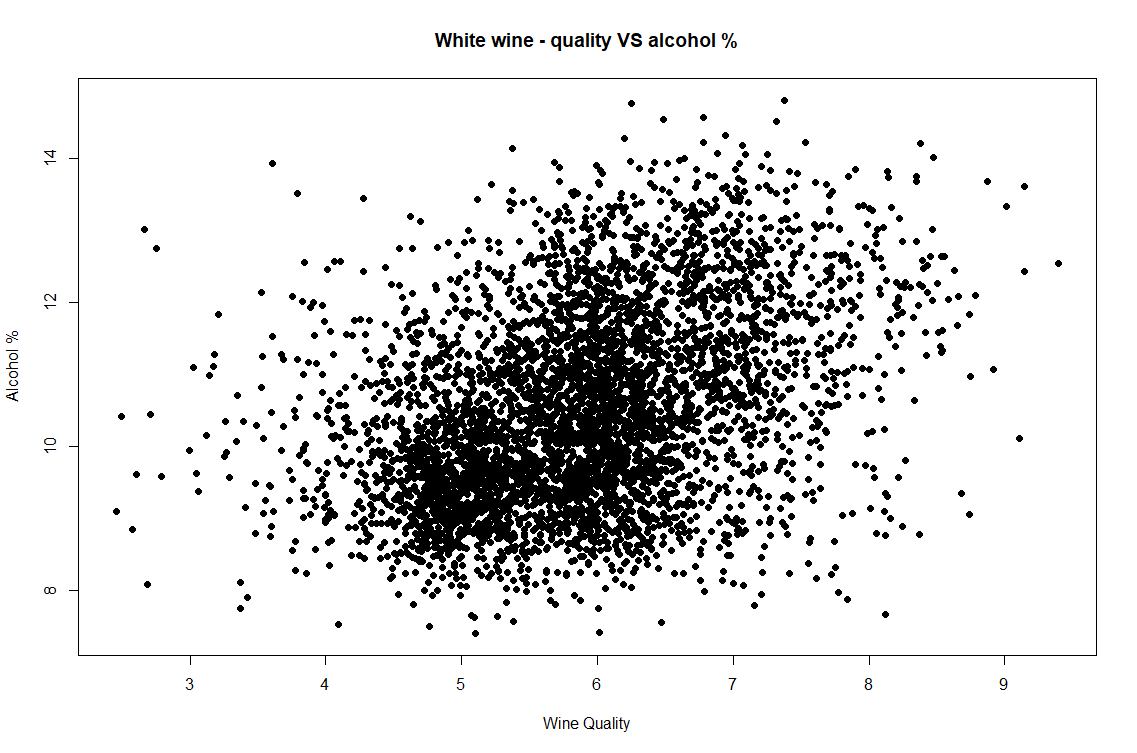
Description automatically generated

Chart

Description automatically generated

1. מעלת הפולינום המתאימה ביותר היא 5, בהתאם לערך SSE הנמוך ביותר. (יש לציין כי ב-overfit רציני, אבל עדיין הכי מתאים).

# שאלה 2

1. 

Chart, scatter chart

Description automatically generated

|  |  |  |
| --- | --- | --- |
| סוג יין | מקדם מתאם | p-value |
| לבן | 0.3896179 | 2.2e-16 |
| אדום | 0.389751 | 2.2e-16 |

נמצא קשר מובהק בין האיכות לאחוז אלכוהול ב-2 היינות, ועוצמת הקשר גם גבוהה בשניהם.

1. ניתן לראות בקובץ ה-R המצורף.
2. ניתן לראות בקובץ ה-R המצורף.

|  |  |  |
| --- | --- | --- |
| סוג יין | 2.5 | 97.5 |
| לבן | -0.0283 | 0.0277 |
| אדום | -0.0494 | 0.0488 |

Chart, histogram

Description automatically generated

Chart, histogram

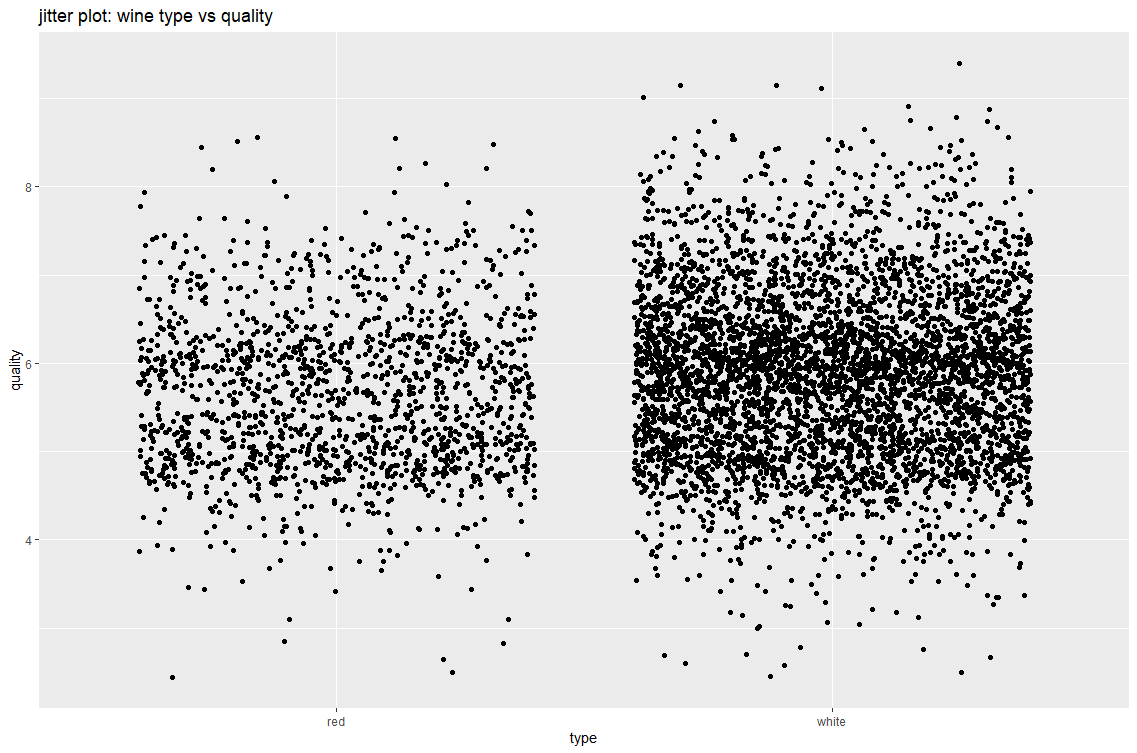
Description automatically generated

1. עבור היין האדום, ייתכן כי זאת בגלל שיש פחות דגימות של יין אדום מלבן.

Graphical user interface, text, application, Teams

Description automatically generated

# שאלה 3

1. 
2. תוצאות המבחן:

U: 3322989   
p-value: 2.2e-16

1. U1: 3322989  
   U2: 4508913
2. Rbc: 0.1514222