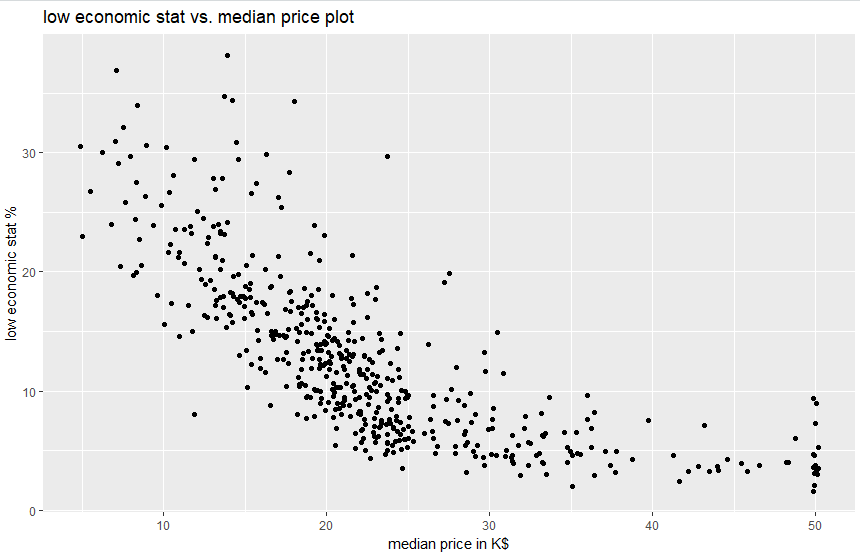
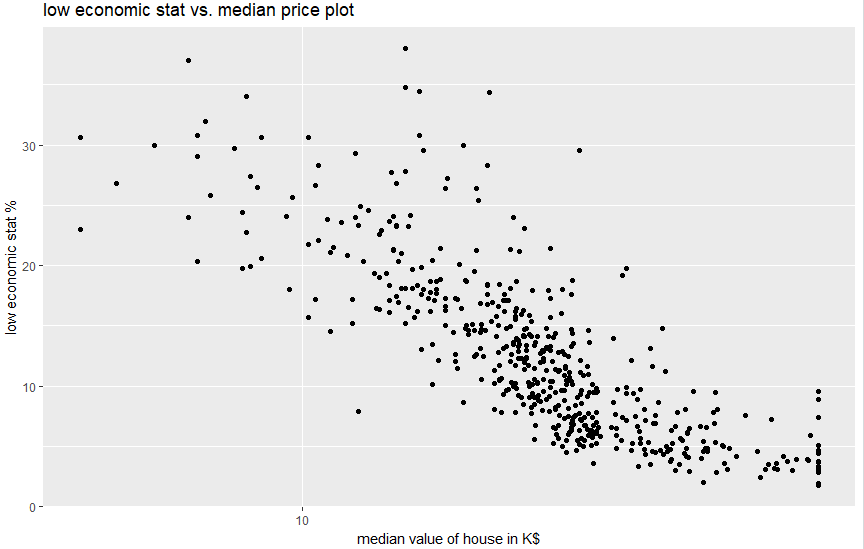
**יישומי R בכריית נתונים, תרגיל2 – קובץ פלט**

1. הקשר בין MEDV ל- LSTAT

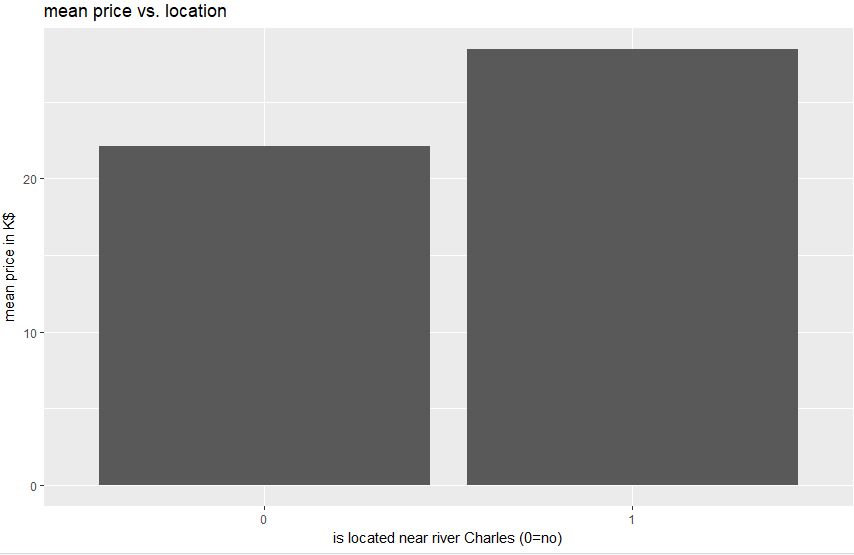
**אנחנו רואים שככל שאחוז האוכלוסייה במצב סוציואקונומי יורד, כך החציון של ערך הבית עולה, הנ"ל נכון עבור ערכים נמוכים של החציון ומתייצב לפלטו באזור ה-25 אלף דולרים.**

ללא הפעלת לוג על ציר X



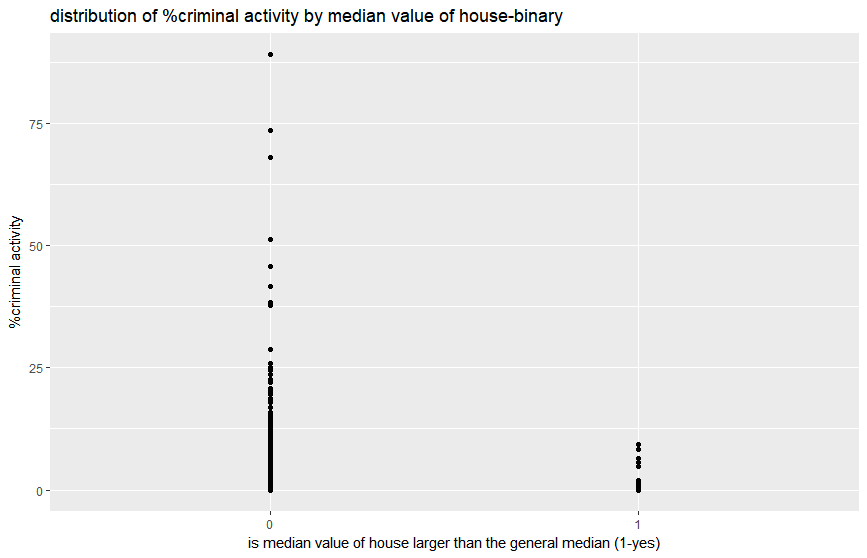
עם הפעלת לוג על ציר ה-X

1. הקשר בין CHAS **לממוצע** MED

**אנו רואים שממוצע ערך הבית גבוה באזורים הממוקמים ליד נהר צ'רלס. האם יש לכך משמעות סטטיסטית? יש להריץ מבחן t להשוואה בין 2 מדגמים ב"ת תלויים**

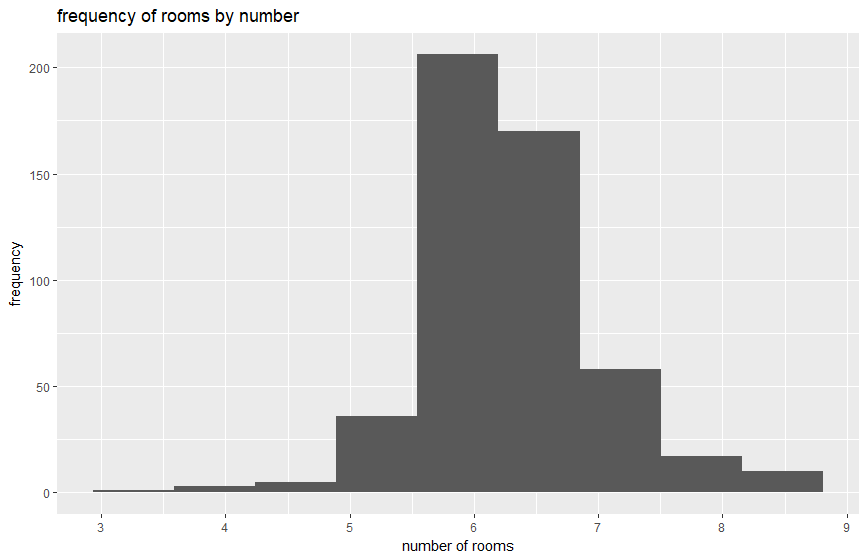
1. התפלגות CRIM לפי MEDV.CAT

**אנו רואים כי במקרה בו חציון ערך הבית נמוך מחציון הבתים הכללי, הפשיעה מתפלגת על פני טווח אחוזים גבוה הרבה יותר מאשר במקרה בו חציון הבית גבוה מהחציון הללי.**



1. גרף נוסף לבחירתך-היסטוגרמה של מספר חדרים בבית

**אנו רואים ששכיחות החדרים נעה סביב הערך 6.**



מאחר ובהגשה ניתן לצרף רק קובץ 1 למודל- מצ"ב קובץ R

