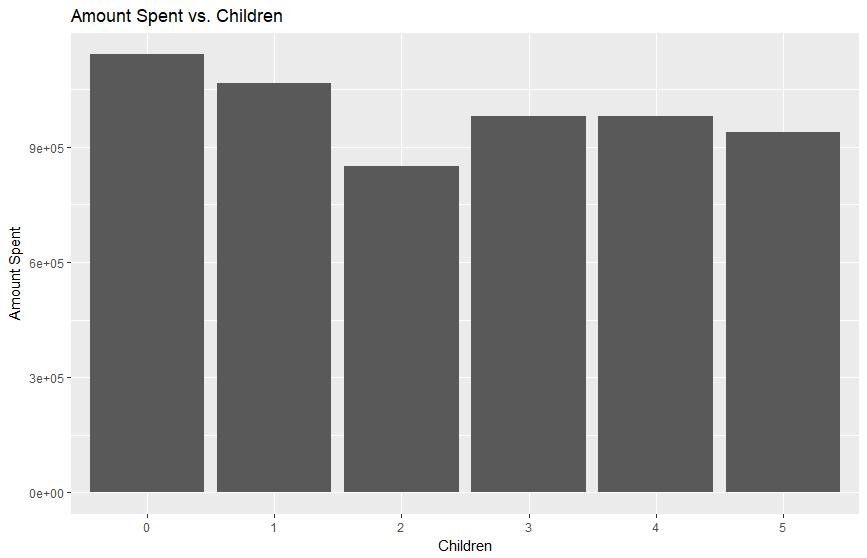
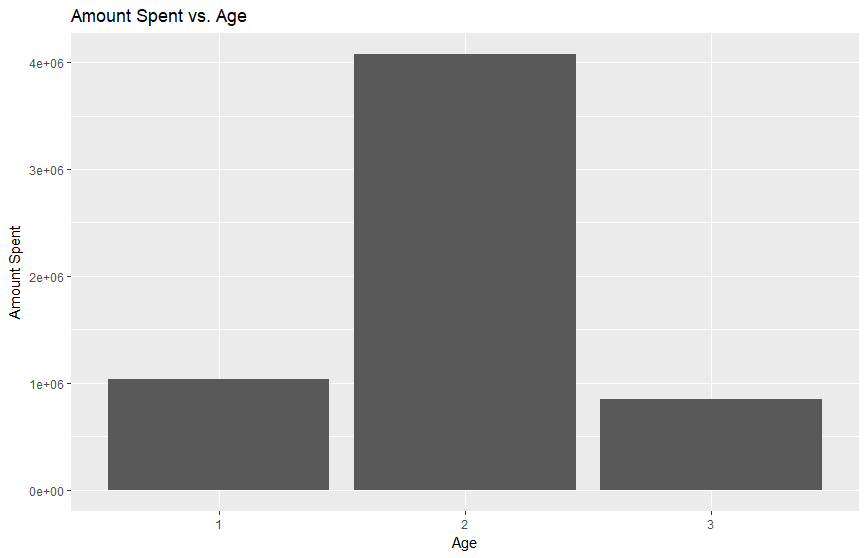
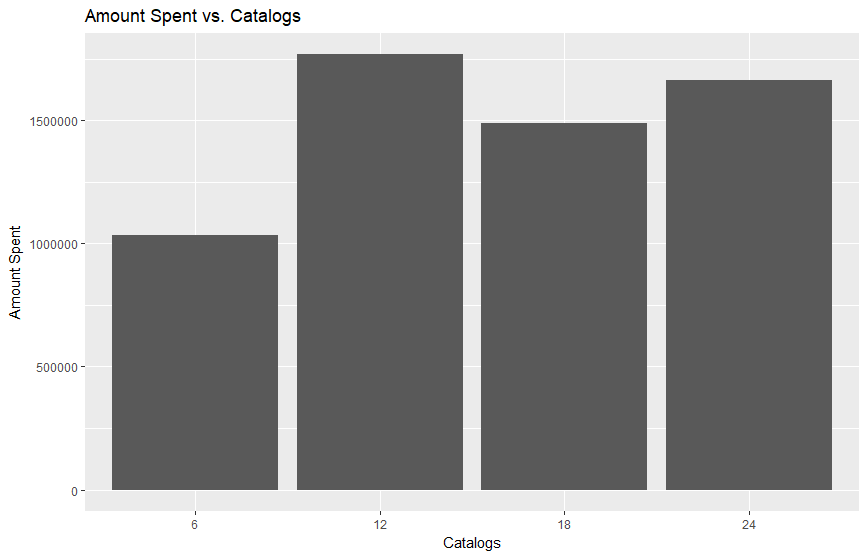
**יישומי R בכריית נתונים, תרגיל 3 – קובץ פלט**

1. הקשר בין המשתנים האורדינאלים למשתנה התלוי-Amount spent:
2. **i. ביחס למשתנה גיל-** צעירים מוציאים יותר ממבוגרים, ואילו אלה שבאמצע מוציאים הרבה יותר משני הקודמים יחד.
3. **ביחס למשתנה ילדים-** לקוחות בעלי 2 ילדים מוציאים פחות מהשאר, לקוחות חסרי ילדים מוציאים יותר מכולם, לקוחות בעלי 3 ו-4 ילדים מוציאים סכום דומה הגבוה יותר מאלו בעלי 2 ילדים ונמוך מאלו חסרי ילדים או בעלי ילד 1.
4. **ביחס למשתנה קטלוגים-** לקוחות להם נשלחו 6 קטלוגים מוציאים פחות מכולם, לקוחות להם נשלחו 12 קטלוגים מוציאים יותר מכולם, לקוחות להם נשלחו 18/24 קטלוגים מוציאים ערכים בין שני הקיצונים מעלה.
5. מאחר והקשר לא לינארי, יש להתייחס למשתנה כאל משנה נומינאלי (הוגדר בתור as.factor).





1. שני המשתנים הכמותיים – AmountSpent ו-Salary מתפלגים לפי התפלגות זנב ימין, רק אחרי שמפעילים לוג (במקרה זה- לוג ע"ב 10) על שניהם, רואים קשר לינארי חיובי בין שניהם. כלומר פונקציית לוג תרמה ל"פתיחת" התפלגות זנב ימין ולהצגת הקשר.



1. 