

Liquor Store Sales Analysis Report:

Παρακάτω αναφέρεται μια πλήρης περιγραφή της εξαγωγής δεδομένων και αποτελεσμάτων, χρησιμοποιώντας δεδομένα από το **finance_liquor_sales** σε Excel και επεξεργασία σε Python με τη χρήση των ενσωματωμένων βιβλιοθηκών **Pandas** για την εισαγωγή και τροποποίηση της βάσης δεδομένων, καθώς και της βιβλιοθήκης αναπαράστασης δεδομένων **Matplotlib**.

Σταδια Επεξεργασίας:

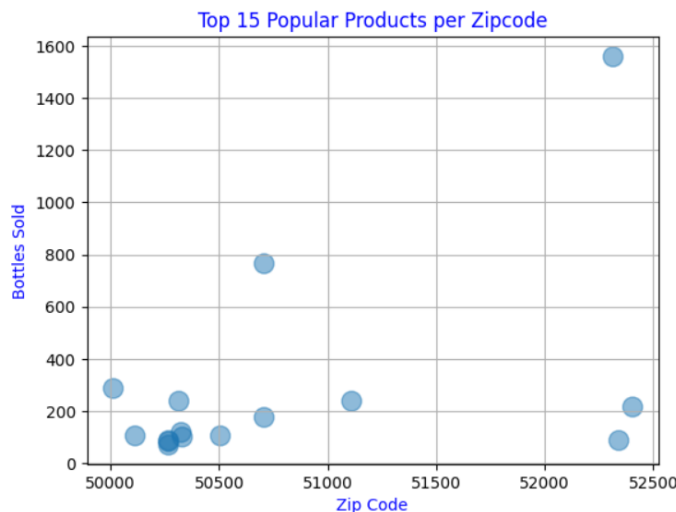
1. Έλεγχος / Κατανόηση / Καθαρισμός Δεδομένων

- Ελέγχθηκε η βάση δεδομένων καθώς και οι στήλες της για κενά στοιχεία. Αρχικά, εξετάστηκαν οι κενές στήλες, και εφόσον δεν επηρέαζαν τα αποτελέσματα, αφαιρέθηκαν.
- Στη συνέχεια, απομονώθηκαν οι στήλες που ήταν απαραίτητες για τους υπολογισμούς, ενώ εφαρμόστηκε πρώτα φιλτράρισμα στα δεδομένα μεταξύ των ετών **2016-2019**.

2. Task A

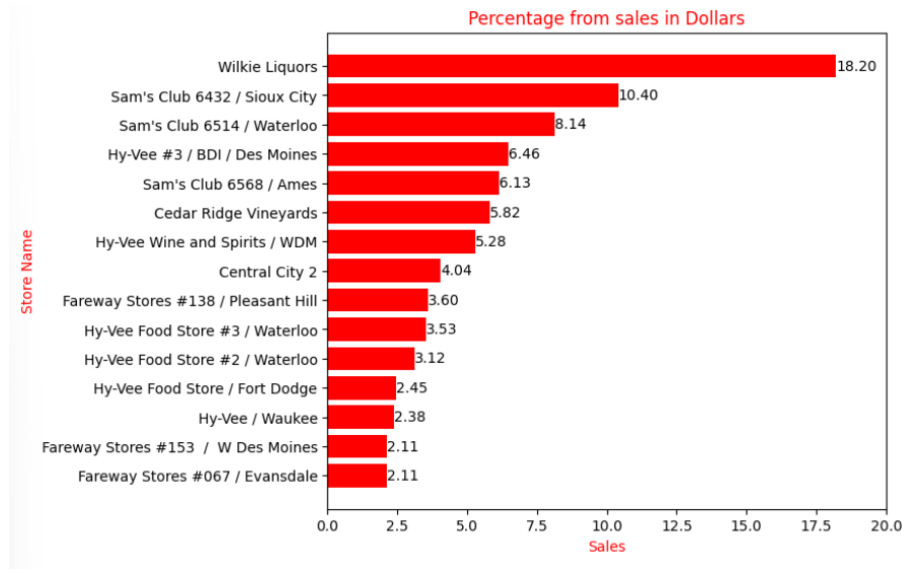
- Για το πρώτο task απομονώθηκαν οι στήλες **Zip_code**, **Item_number**, **bottles_sold** και εκτελέστηκαν οι κατάλληλες εντολές ώστε να βρεθούν τα δημοφιλέστερα προϊόντα ανά ταχυδρομικό κώδικα.
- Τα αποτελέσματα παρουσιάστηκαν σε **διάγραμμα διασποράς (Scatterplot)**, αναδεικνύοντας ότι το πιο δημοφιλές προϊόν καταγράφηκε στον ταχυδρομικό κώδικα **52314**, όπου πωλήθηκαν **1.560** προϊόντα.

Παρακάτω το διάγραμμα :



3. Task B

- Για το δεύτερο task απομονώθηκαν οι στήλες **store_name** και **sale_dollars** και εκτελέστηκαν οι κατάλληλες εντολές ώστε να υπολογιστεί το συνολικό κέρδος από όλα τα καταστήματα.
- Στη συνέχεια, υπολογίστηκε το ποσοστό συμμετοχής κάθε καταστήματος στο συνολικό ποσό, το οποίο παρουσιάστηκε με **οριζόντιο ραβδόγραμμα (Horizontal Bar Chart)**.
- Αναδείχθηκε ότι το κατάστημα με το μεγαλύτερο ποσοστό σε πωλήσεις (\$) ήταν το **Wilkie Liquors** με ποσοστό **18,20%**, όπως φαίνεται και στο παρακάτω διάγραμμα:



Βλαδιμηρος Ποζιδης