Data Intro The dataset consists of information about the purchases of chocolate candy bars of 500 individuals from a given area when entering a 730 41K 7151M physical 'FMCG' store in the period of 2 years. All data has been collected through the loyalty cards they use at checkout. Sum of Quantity ·Identify the key factors that are impacting the purchases of chocolate candy bars. Work on the data and generate actionable insights. Count of Gender by Gender Count of marital status by Count of City by City Count of Employee by Quantity and marital status Sum of Quantity marital status marital status onon single single 32K 30K 5 20K 23K 23K 40K 15K 20K 20K 12K 10K small city mid-sized big city 0K single City Count of Educations by Educations Count of Employee by Employee Sum of Income 21K 30K € 20K **□** 20K 8K 10K 7151M Quantity skilled **14303M** Sum of Quantity by Educations and Gender **OM** employee Gender • female • male **Using the Demographic Data about people EDA** Insights we are seeking to get out of the analysis is to understand the purchasing patterns of the 20K segments, Which would Answer Questions such as Quantity 15K ·Which Segment has more monetary value? ·Which Segment Purchases more often? ·Which Brands do each segment prefer more often? 10K ·Who are the most valuable customers(Overall and to each brand) and which segment do they belong 5K to? 10K Sum of Income by Employee Sum of Income by Educations 5K 5K 0K high school graduate school 1bn university unknown 1bn (12.34%)(2.7%) Educations (19.93%)Count of ID by ID and incidence incidence onot purchased purchased **Educations Employee** high school university 10 unemployed unknown 11 11 highly qualified employee graduate school 2bn (24.85%) 50 98 92 87 2bn 4bn 76 74 (53.3%) (26.77%) 4bn (60.11%) 20000001 200000002 20000003 20000004 200000005 Sum of Quantity by Brand Count of ID by City ID 14K 300 Sum of Quantity by Age and Brand 12.8K **Brand** •0 •1 •2 •3 •4 •5 280 12K 11.2K 2000 10K 200 8.9K Count of ID 8K Quantity 121 6K 99 100 4.7K 4K 3.0K 500 2K 0.0K small city mid-sized big city 0K 2 3 5 **city** City 0 4 Brand Age Count of ID by City and Gender Sum of Quantity by Age and marital status Sum of Quantity by Employee and Gender Gender ● female ● male marital status • non single • single Gender • female • male 300 250 20K 152 200 14K Count of ID Sum of Quantity 10K 7K 100 74 8K 3K 128 80 5K 50 2K 0K 47 highly qualified skilled employee unemployed 0 employee big city small city mid-sized Employee, Age and Income 60 **Employee** • highly qualified employee • skilled employee • unemployed **Observations** · As we can see above Loyal Customers actually bough the most candy 300K · Most number of people are in Potential Loyal Customer, It's always good to know that there is so much potential! 250K Except for brand 2 in which 'customer needing more attention' segment has highest purchases, among all the other brands champions purchased more · cross elasticity value of Brand-1 corresponding to 1.39 indicate the factor by which our brand's(brand-2) purchase 150K probability could decrease, so a decrease by 1% of brand-1 price would mean a decline of 0.2% purchase probability of brand-2 · We could counter that discount given by brand-1 by decreasing our brand's price by 1% in which case purchase 100K probability of our brand would increase by 1.2%

50K

· Which means as a strategic move to counter the 1% discount of price by brand-1 we discount our brand candy price

· If we wanted to find out by how much do we need to decrease our prices in order to counter the decrease of brand-1 then we can simply do (cross_elasticity/brand_elasticity), which is (0.2/1.2 = 0.16) so we need to discount our prices

by 1% then total we would have an increase of 1% purchase probability (as 1.2-0.2=1)

by 0.16% to retain the market share