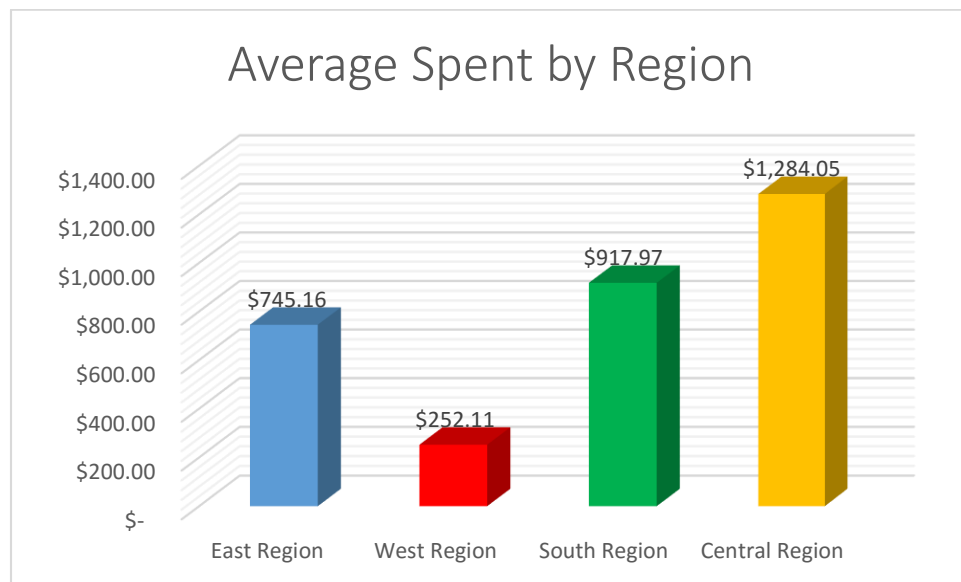


## Customer Buying Patterns Report

The key to a successful business is understanding the customer and their needs. Staying current with new trends in business and in customer spending is key to that success. Through an in-depth analysis of the customer transaction data combined with a variety of data mining methods, we are able to come up with many degrees of insight into customer buying patterns. The main goal is to see what factors are relevant in determining customer purchasing behavior and then hopefully using that data to predict shopping habits from different key demographics, including age, region, and if the customer purchased products online or in-store.

Before we get started with the investigation of customer buying patterns it is important to gather basic conclusions drawn from the data. From the 80,000 lines of data given to me I felt the first important task was to figure out the average purchase separated by region. The average spent

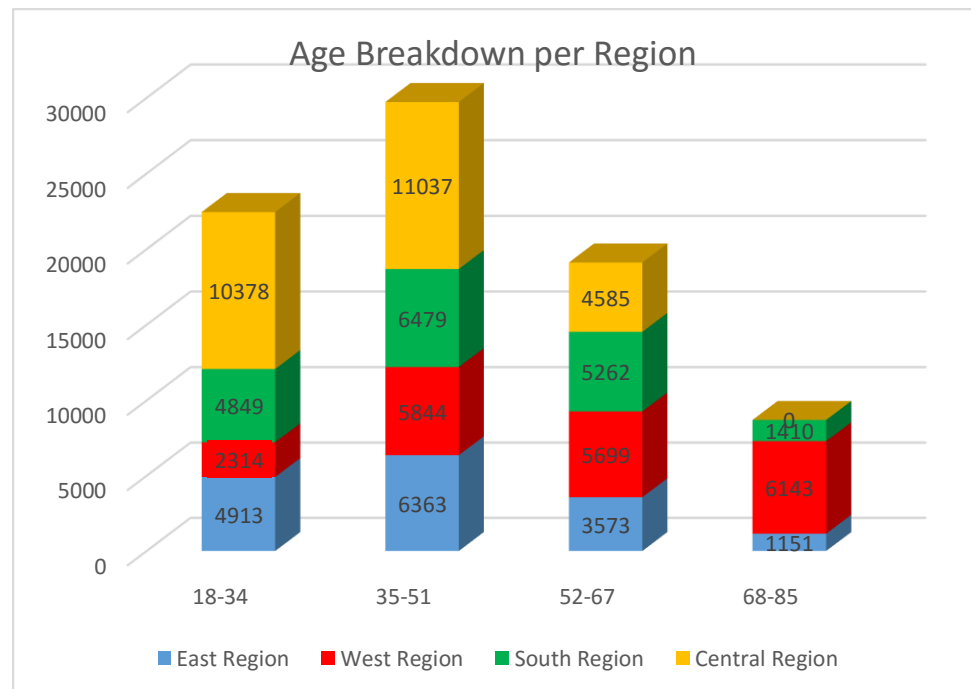


for all complete transactions comes to \$835.92 but the important things to take away is of the four regions, the central region has tendencies to spend over \$350.00 more per average than the south

region, which came in second. Of the four regions the west is clearly the worst performer, with \$252.11 and having over a thousand-dollar difference per average when compared to the central region.

Our next concern was trying to understand the differences in the age of the customers between regions and then from that data, attempting to predict the age of the customer and their spending habits based on region. In order to attempt to predict the age I separated the customers into 4 different age groups so we could categorize certain shopping habits in a generational standard. These groups are

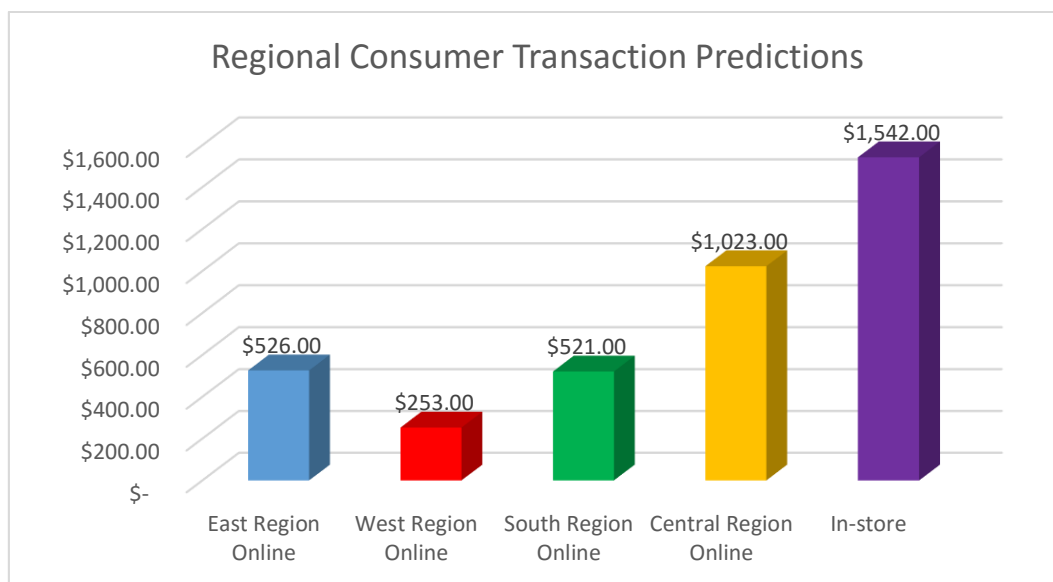
18-34, 35-51, 52-68, and 68-85. Now based on the general data alone, we were able to infer certain things about the age of customers between regions. The most notable was that not all age groups were



represented in all regions. The largest region, the central region, was predominately younger, with 18-34 and 35-51 doing most of the purchasing in that region while the 68-85 year olds were not represented at all. The west region had the most 68-85 customers and the least amount of 18-34 year olds. The east and south regions were both represented by all four age groups in an almost equal balance. Even with all this information and data at hand, predicting the age of the customer by region proved to be inconclusive. After analyzing the data, the best rate of prediction was with a 38% accuracy. This is too low to confidently predict the age of the customer by region. Perhaps with more information

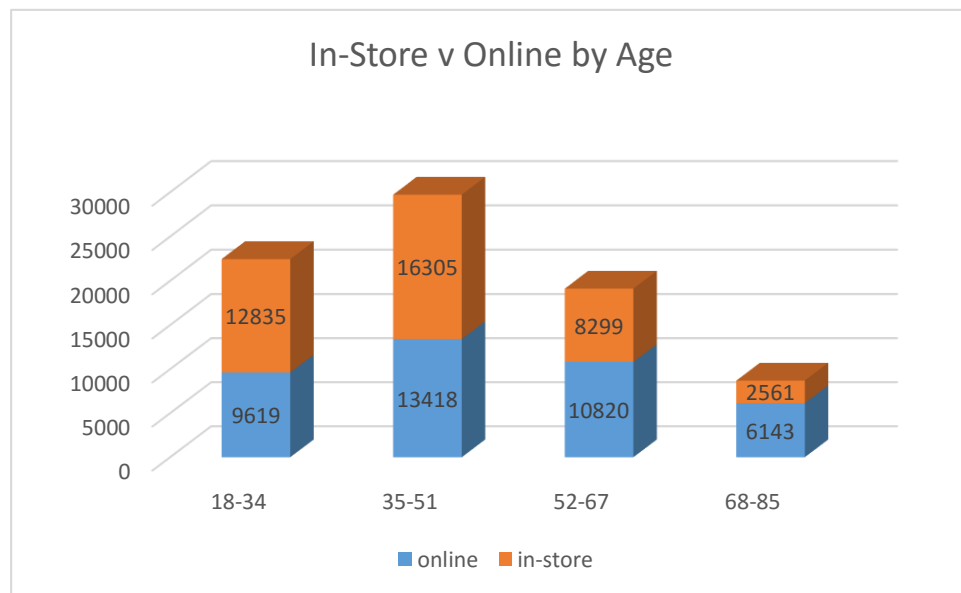
one could get a better percentage but as of now correctly predicting the age of the customer is inconclusive.

Our next task was to determine whether or not we can use the data that was given to predict the amount a customer will spend per transaction. By using the data concerning region, age, and in-store or online purchase we were able to predict the amount a customer will spend per transaction by region online only. Online shopping in the central region was the highest, at \$1,023, and the west was the lowest, at \$253. In-store predictions were not separated by region but rather grouped together as a whole, and came in at \$1542. The west region shopped exclusively online so they were excluded from that side of the report.



Our last major concern deals with online shopping versus in-store and how that correlates with age. There was a 90% success rate in predicting the age based on shopping habits. The main factors that helped realize such a high percentage, besides age obviously, were region and amount spent per transaction. Considering the east region did all their shopping online and the west all in-store, compared with the west region having the older customer base and the central region with the

youngest, this combination of data compared with the average amount spent per age group was pivotal in the ability to predict if customers will buy online or in-store.



Overall, we discovered that through different data mining methods, we are able to come up with a variety of insight concerning customer buying patterns. Although not every task was a complete success, much can still learned from the data. Understanding spending habits by age, region, and method will help Blackwell understand their customer base and in turn help set up the corporation for success. Although we were unable to predict the age of a customer based on region, I believe it possible to predict such things, but different or more detailed data would be necessary. Luckily, we were able to succeed in figuring out some of the tasks given to us. The amount a customer spends per transaction is valuable insight and with proper marketing enhancements should prove to be invaluable to the growth of Blackwell. Likewise, being able to successfully predict if a customer is making purchases online versus in-store gives Blackwell an advantage in marketing techniques and also should help the company immensely. The data mining methods used in this report will definitely aid in the continued success of Blackwell and their eCommerce Team now and into the future.