T1 questions:

- Do customers in different regions spend more per transaction? Which regions spend the most/least?

- Is there a relationship between the number of items purchased and amount spent?

T2 questions:

- The correlation between the age of a customer and where the transaction took place

- The correlation between number of items purchased and amount spent

Methods used:

* EDA
* ML algorithms related to Classification

Results delivered:

How can data mining be used to make data-driven decisions for all aspects of a business?

* Identify regions that are least profitable and most profitable
* Identify potential trends and relationships in data
* Develop tools that can predict different aspects of a customer based on historical data about a group of customers
* Develop informative graphs that can answer key-business questions
* Provide stakeholders will knowledge they need to optimize or expand their business-related tasks

Opportunities/Insights 1 ---

* Online shopping is as popular as retail shopping.
* Show matrix(s)/visualization(s) of shopping preference based on age, region, and amount spent

Opportunities/Insights 2 ---

* The Southern region has potential to grow. Customers in that region should be given additional incentives to encourage spending, such as savings and generally lower prices.
* Show matrix(s)/visualization(s) that showcase that the Southern region tends to spend less, and how this should be interpreted as an opportunity to sell more discounted items to drive up revenue in the region
* Show histogram of amount by region and highlight the East region.

Opportunities/Insights 3 ---

* Multi-item purchases make up the overwhelming majority of purchases. So, Market Basket Analysis should be used to generate more effective product groupings in order to encourage customers to spend more. This can also be coupled with discounts and other promotional deals in order to drive up revenue.
* Show histogram of item amounts to highlight that multi-item purchases are exceedingly common versus single-item purchases.

Opportunities/Insights 4 ---

* The North and East regions have similar total revenues which may mean that marketing strategies in those regions need not differ too much. Purchasing patterns in the two regions may be similar enough to justify not having to develop and deploy completely different marketing strategies in those regions, which can reduce overall expenses.
* Show histogram of amount by region and highlight how similar the North and East regions are.

Opportunities/Insights 5 ---

* The West region is the most profitable region for the company and should continue to be marketed so as to encourage continual customer spending. That said, the company should not put “all of its eggs in this basket”. As important as the West region is for profitability, it does not make sense to exclude the other regions.
* Show histogram of amount by region and highlight the West region.

Opportunities/Insights 6 ---

* About 25% of total purchases have a cost of $1200 or more, which means that Blackwell should not try to bring in more “whales”. Instead, more efforts should be done in expanding to potential customers who spend between the ranges of $50 and $400 per transaction, as this range of spending accounts for the remaining 75% of total purchase spending.
* Show the describe() matrix.

Opportunities/Insights 7 ---

* Blackwell should continue to market its products and services for the working age population, those who are 20-60 years old, as this age cohort spends the most per purchase and most frequently. Blackwell should not prioritize elderly (65+) and young (18) customers, as they appear to be with less disposable income.
* Show histogram of ages and age cohorts.