

Data science case 2: questions

Introduction

This document lists a short set of questions to determine in what way your approach has separated your end result from that of the other teams. In this sense, these questions should not be viewed as a test of the quality of your work, rather illustrate how different approaches lead to different results.

Questions

Which store has the highest sales across your entire date range?

Which item family was sold the most in West-Bengal?

Is there a relationship between oil price and sales across all stores? How did you decide this?

Do you find a connection between the amount of sales and the days of the week? Do weekend sell more or less? What about holidays?

Did you notice negative sales? How did you handle these in your dataset?

Which type of store is most present in your data?

Besides West-Bengal, which state has the highest amount of stores?

Does the items data contain more perishable or non-perishable items?

What is the most selling item across all stores, all dates?