

Northwind DB Hypothesis Testing





Hi there! I'm a software engineer based in NY and working for the UN, and an aspiring **data scientist** at Flatiron Academy

Barto Molina, Flatiron DS, April 2019 cohort



| the company



77 products from 29 suppliers



8 categories



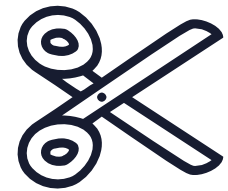
830 orders shipped by 3 shipping companies



9 employees covering 53 territories in 4 regions



91 customers from all over the world



discount **product quantity**

Does discount amount have a statistically significant effect on the quantity of a product in an order? If so, at what level(s) of discount?



customer region **order price**

Does the customer region have a statistically significant effect on the total price of an order?



employee **order price**

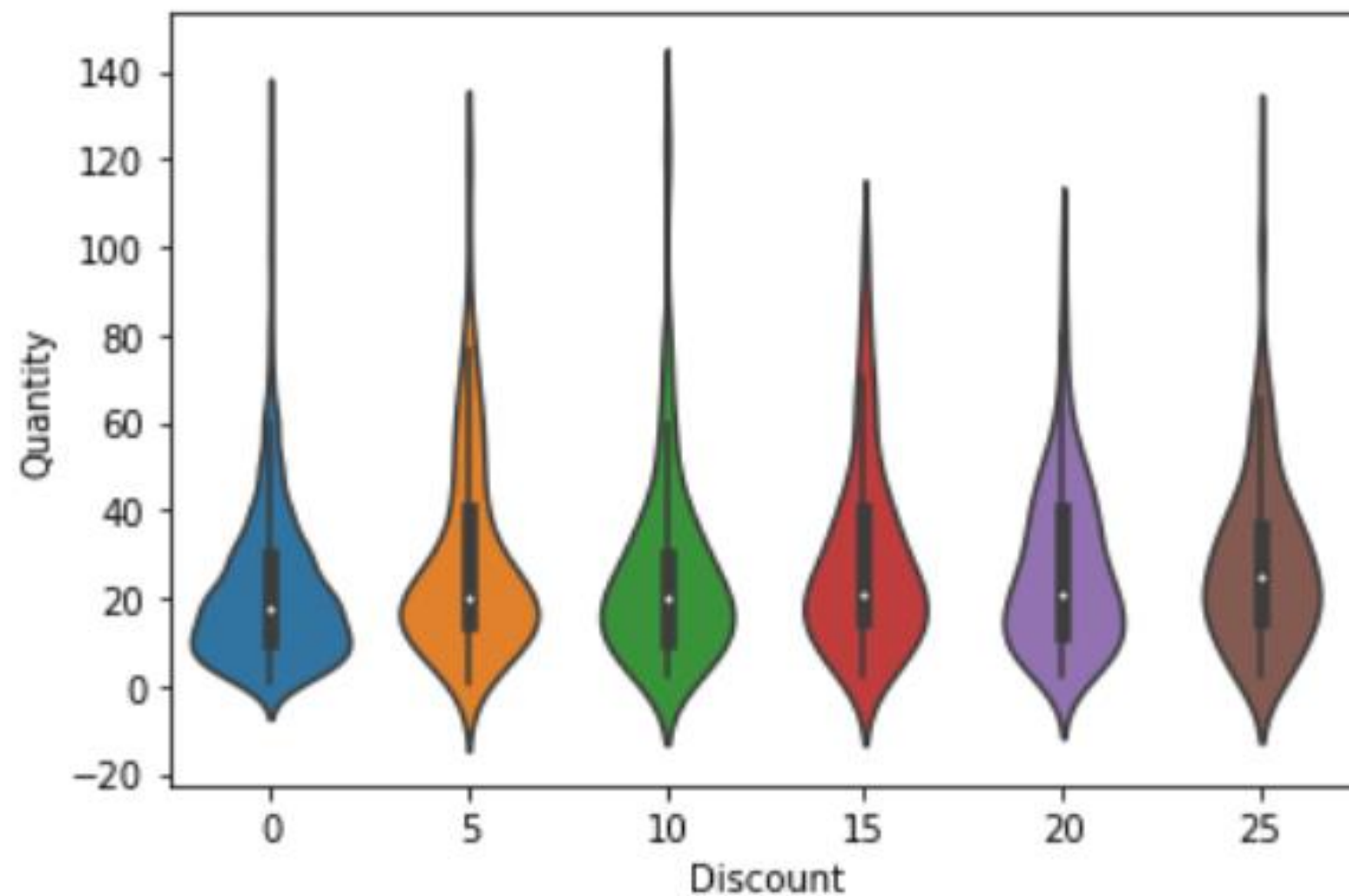
Does the employee have a statistically significant effect on the total price of an order?

discount | product quantity

Main discount levels

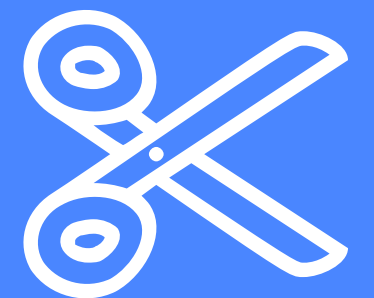
0%, 5%, 10%, 15%, 20%, 25%

Group orders by Order / Product



15% highest effect on the product quantity

Welch's T-test: **5.02 e-11**
Effect Size: **0.3** (15% - 0.37)



Discount



customer region
order price

Customers from 9 different regions

Total price of the order by customer region

Group orders by Customer / Order



region

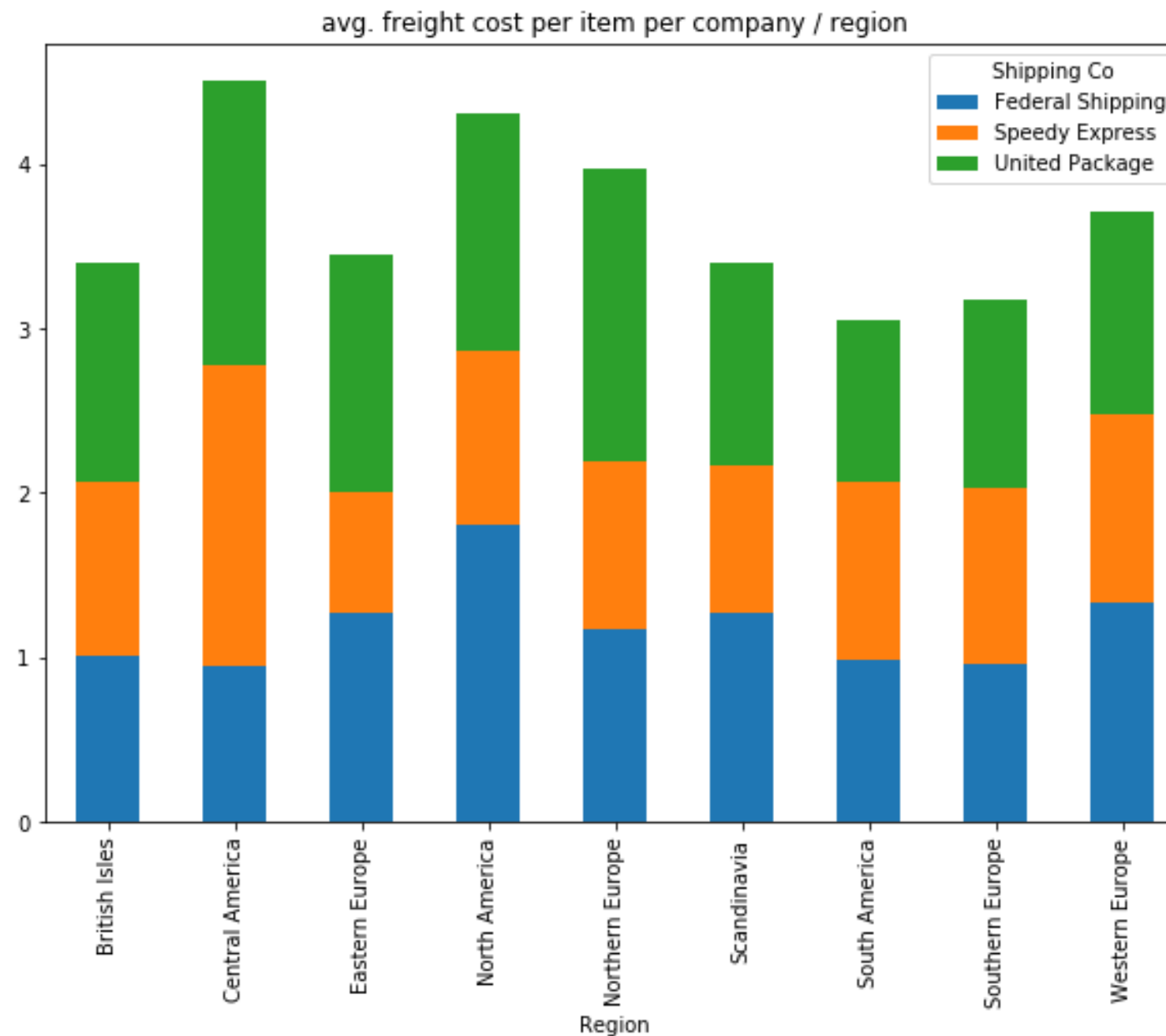




region

customer region
order price

Does freight make a difference?



We can't conclude that the freight cost has an influence on the order price

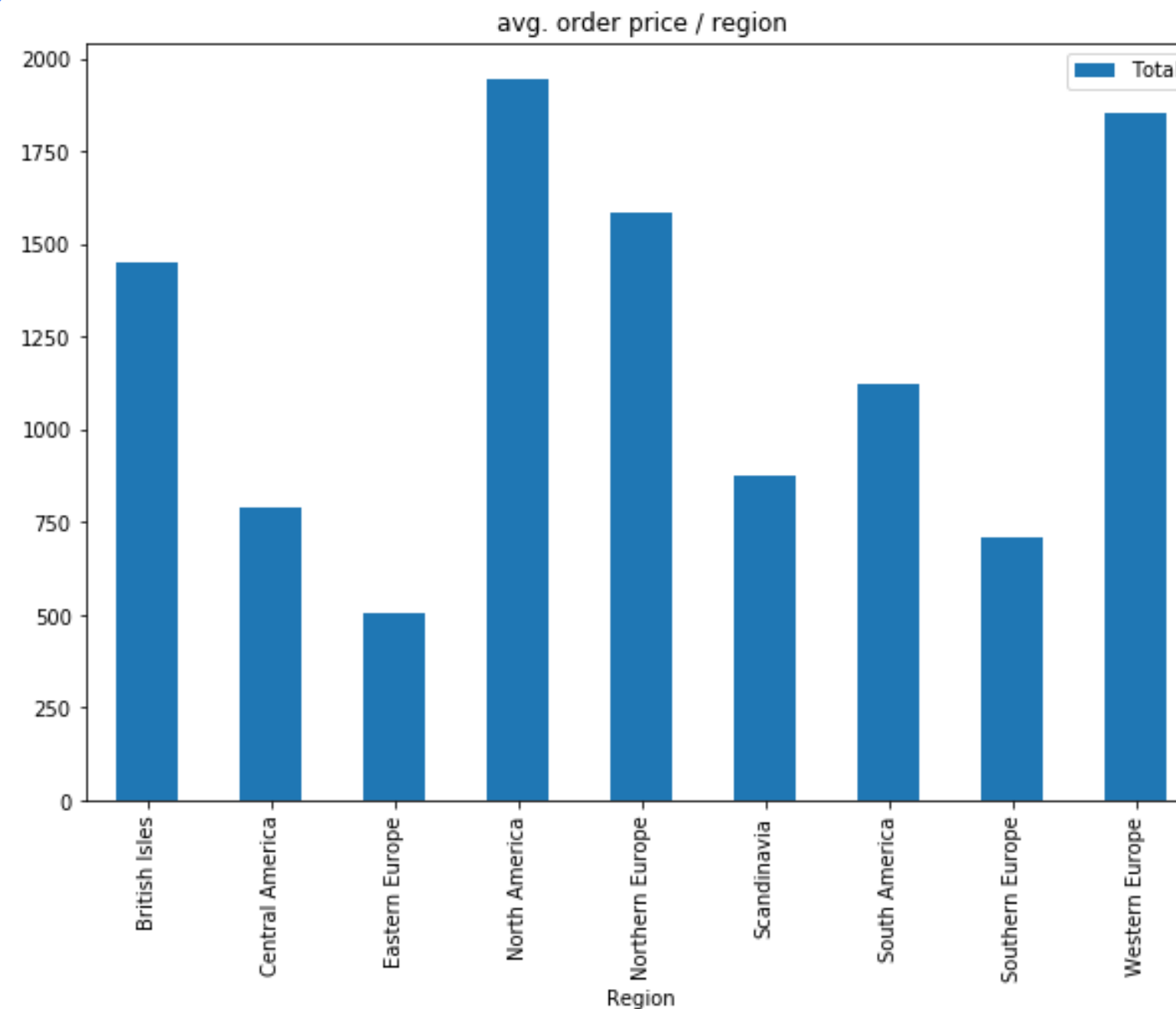




region

customer region
order price

Price by region



The region **has an impact** on the total price of the order



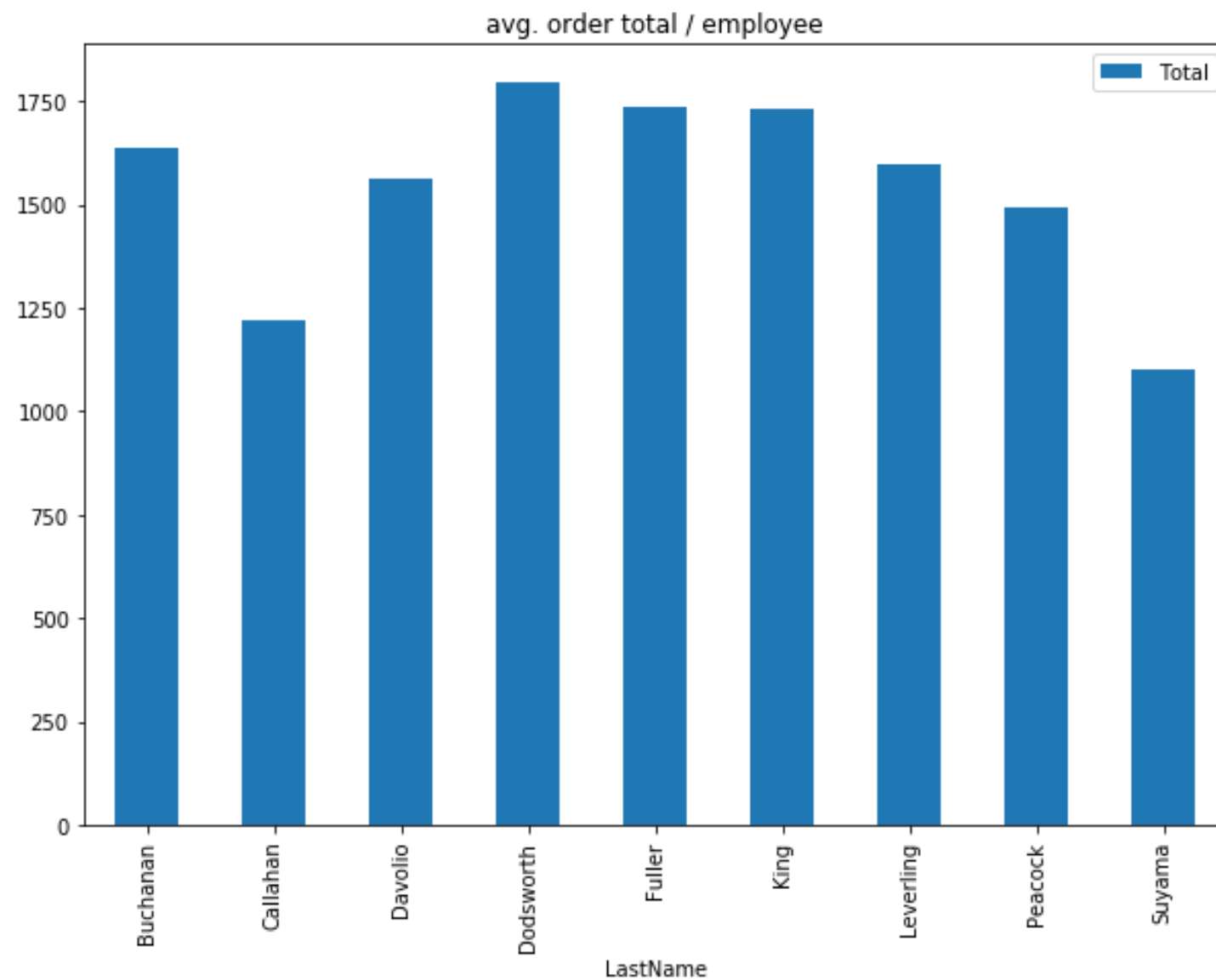
employee
order price

9 employees

Total price of the order by employee



employee



Dodsworth: \$1,798

Suyama: \$1,103

We **can't conclude** that the employee has an impact on the total price of the order



final recommendations



Max. 15% discount



Focus Southern / Eastern Europe



Review employee compensation scheme

「thank you.」