Wolt Data Analyst - Assignment

Alexandra Neagu

2023-01-09

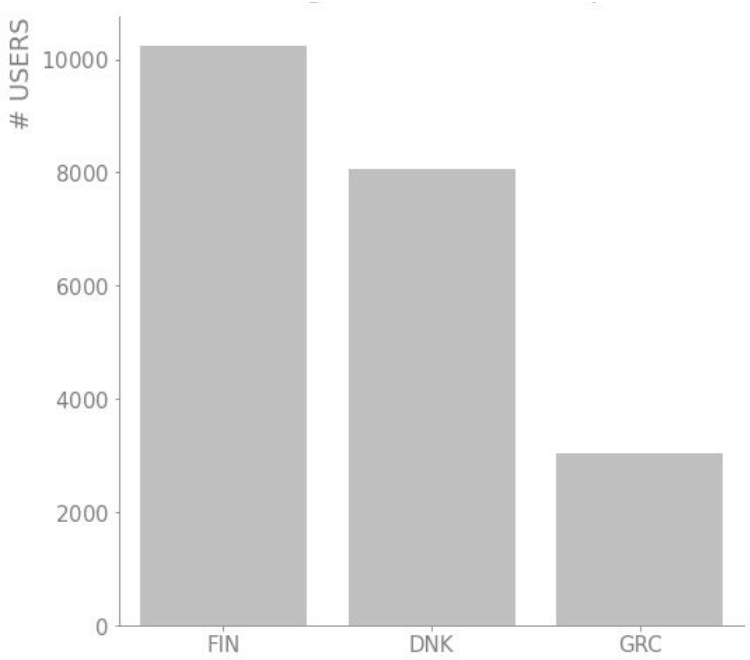
Outline

1. General explorative analysis
2. Active users
3. User segmentation
4. Comments
5. Appendix

General explorative analysis

From **21.9K** users that registered in September 2019 **97%** were from Finland, Denmark and Greece.

Number of registered users for top 3 registration countries

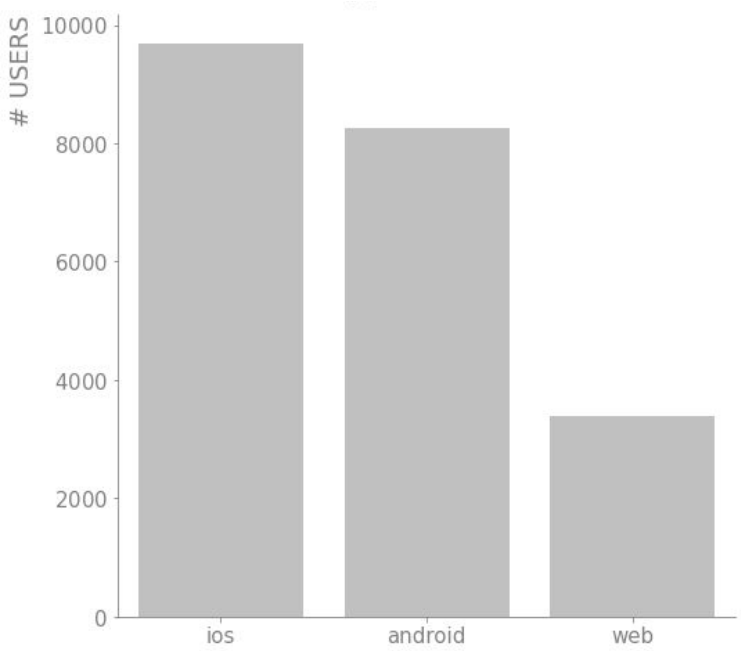


- As Wolt* is a very localized business we do not have enough data to draw any conclusions about the other **56** countries that together account for the rest **3%** of registered users.
- Hence, the analysis will focus only on Finland, Denmark and Greece where in total there were **21.4K** users.

*In this presentation I will refer at the data as being user data for Wolt

84% of users prefer mobile devices to engage with Wolt and **ios** is taking the lead.

Number of registered users by preferred device

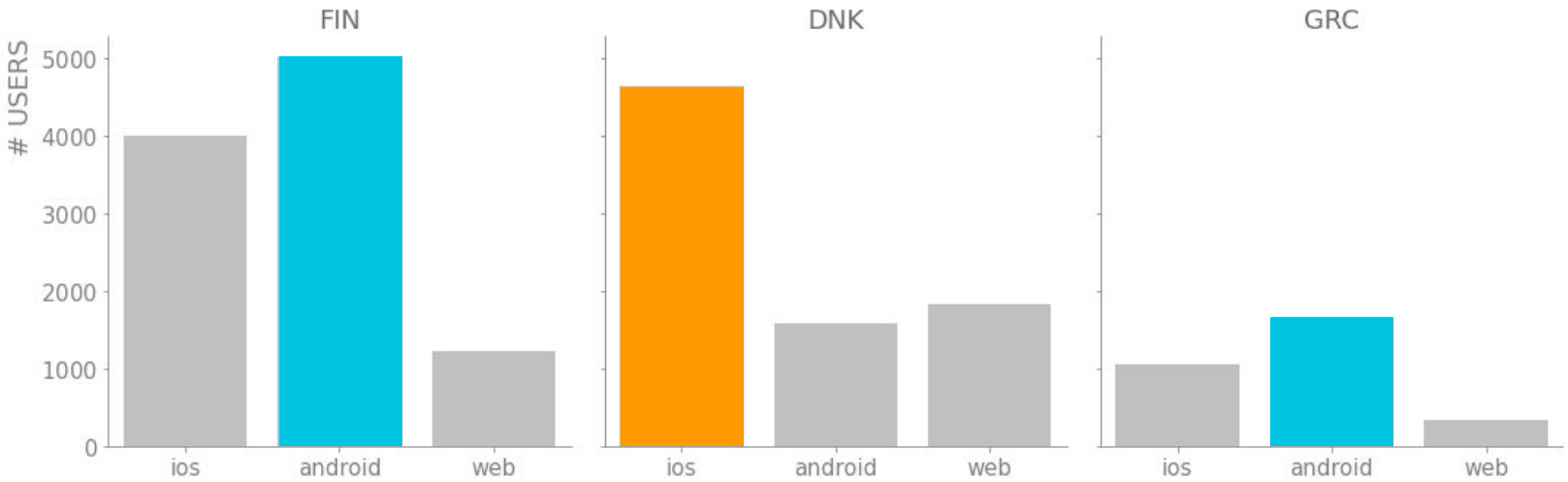


- ios is preferred by 45% of users, android by 39% and web by 16%.
- For users that made no purchase the preferred device is the device they used to register, while for users that made at least one purchase the preferred device is the one used to make most purchases.
- Users with missing data for preferred device were excluded from the analysis.

*see Appendix for more detail on data wrangling for preferred device column

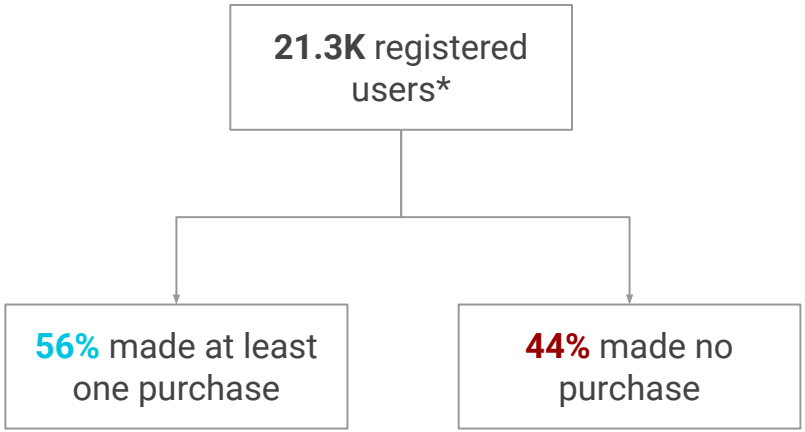
On a country level only in Denmark **ios** is the preferred device while in Finland and Greece **android** is most preferred.

Number of registered users by registration country and preferred device



- On a country level most users prefer to engage with Wolt using a mobile device with a small percentage of users preferring the web.

44% of registered users made no purchase with Wolt within 1 year after registering.



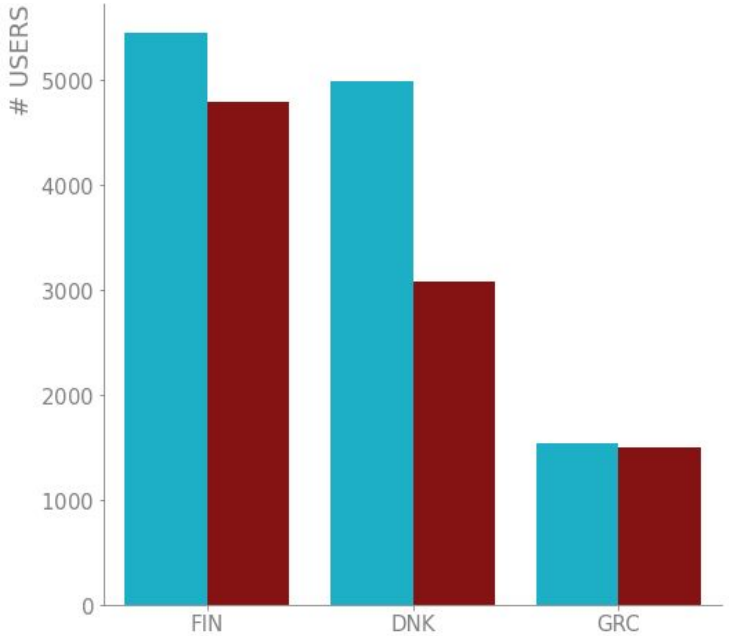
Active users = users that made at least one purchase

Inactive users = users that made no purchase

*number of users for Finland, Denmark, Greece after removing users with no data for preferred device

Distribution of **active users** and **inactive users** is relatively even per country, with Denmark sticking out with more active users than inactive.

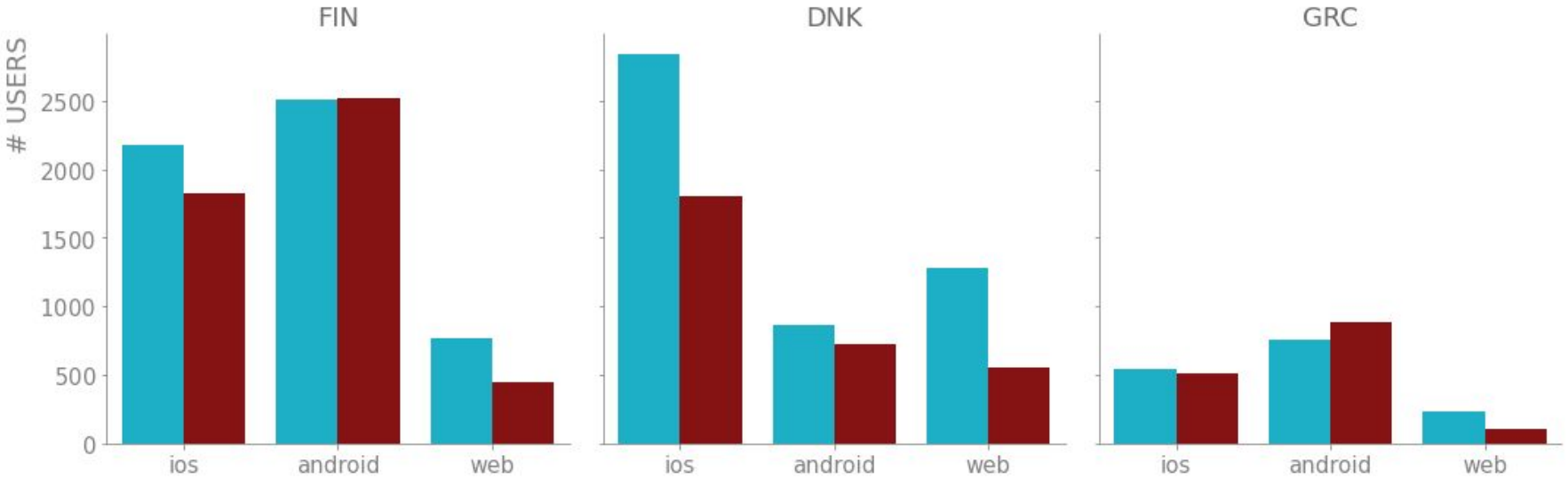
Number of **active users** and **inactive users** by registration country



- Finland and Greece have similar number of **active** and **inactive users** with **53% active users** for Finland and **51%** for Greece.
- Denmark has a bit more **active users** amounting to **62%**.

Distribution of **active users** and **inactive users** for each country is similar between devices.

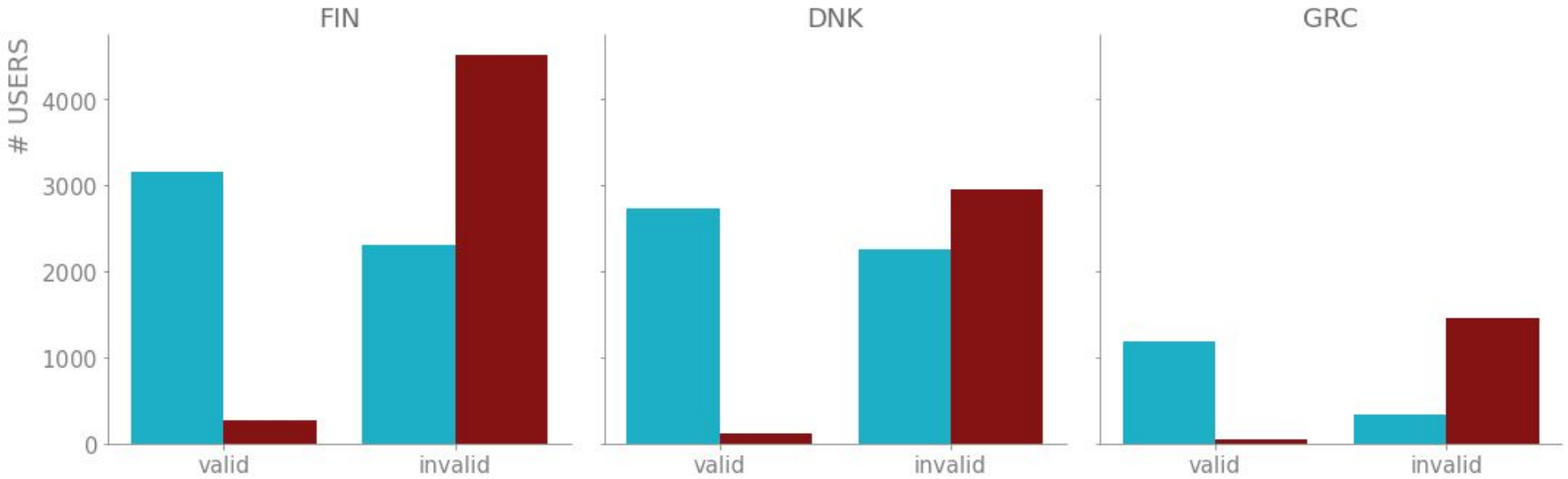
Number of **active users** and **inactive users** by registration country and preferred device



- We do not have a dominant device where users register and remain inactive.

A larger share of **inactive users** have not setup a valid payment method when compared to **active users**.

Number of **active users** and **inactive users** by registration country and payment validity

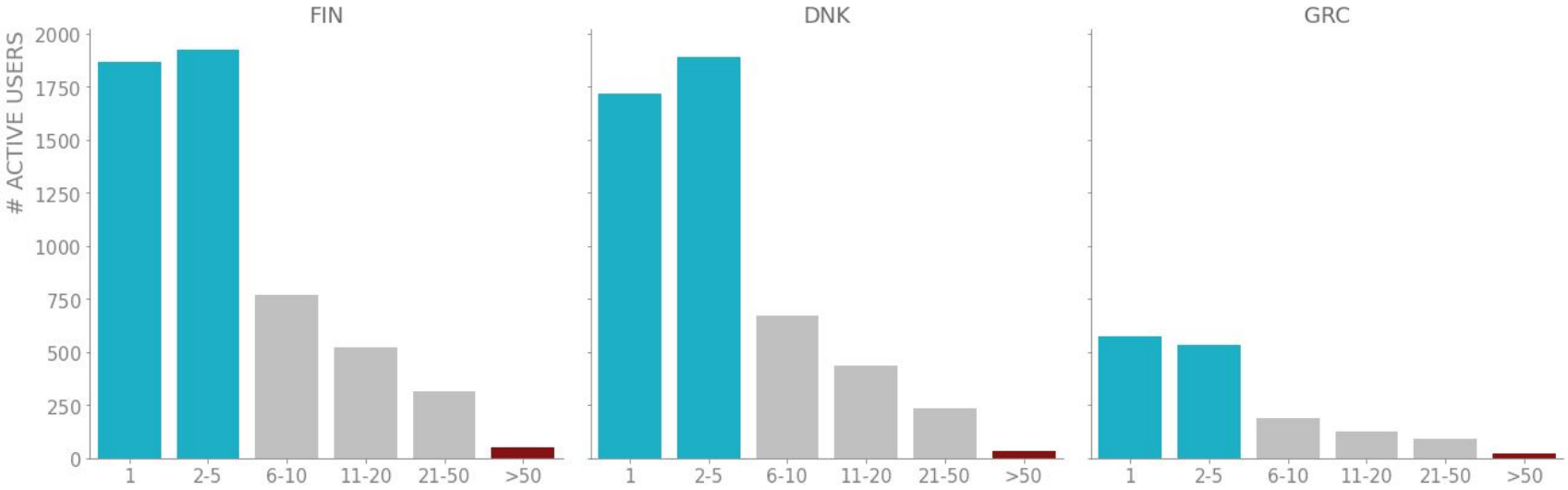


- As expected **inactive users** that have never made a purchase usually do not have a valid payment method setup.

Active users

Around **70%** of active users in each country made either **1** or **2-5 purchases** within one year after registering to use Wolt in September 2019.

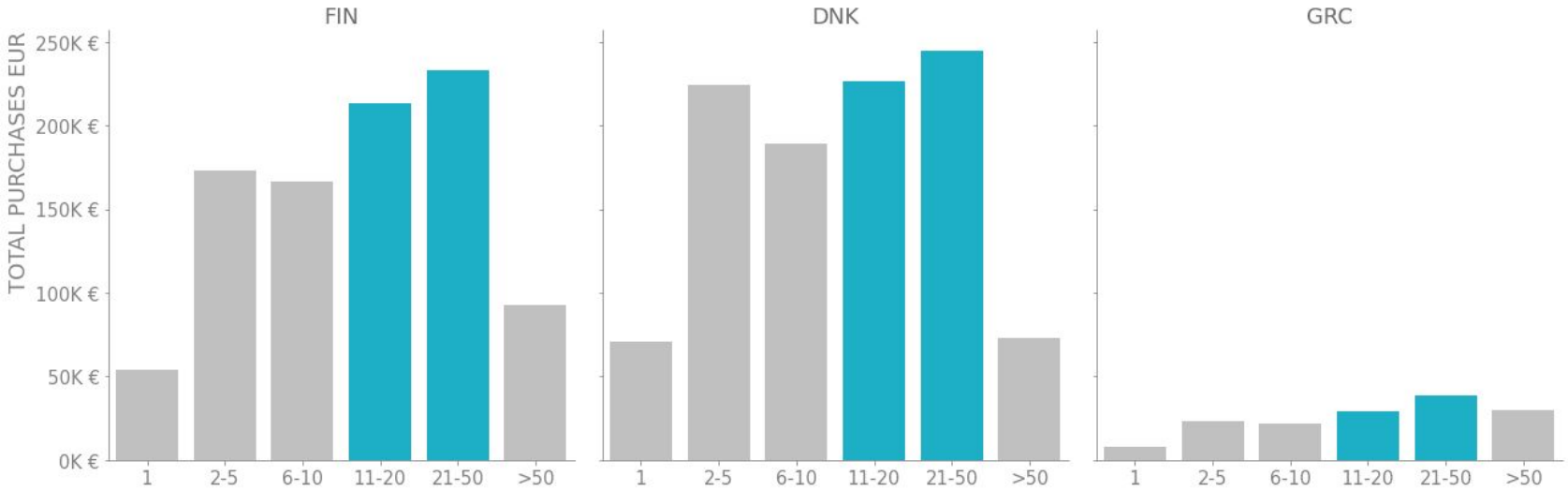
Number of active users by registration country and groups for number of purchases



- If we aim for the users to make at least one purchase per week we would have ~52 purchases within one year. However, only **107 users** in all 3 countries have made **>50 purchases**.

Most money is spent by users making **11-20** and **21-50 purchases** which account for **~1M EUR** and **47%** of the total purchase value of all active users in the 3 countries.

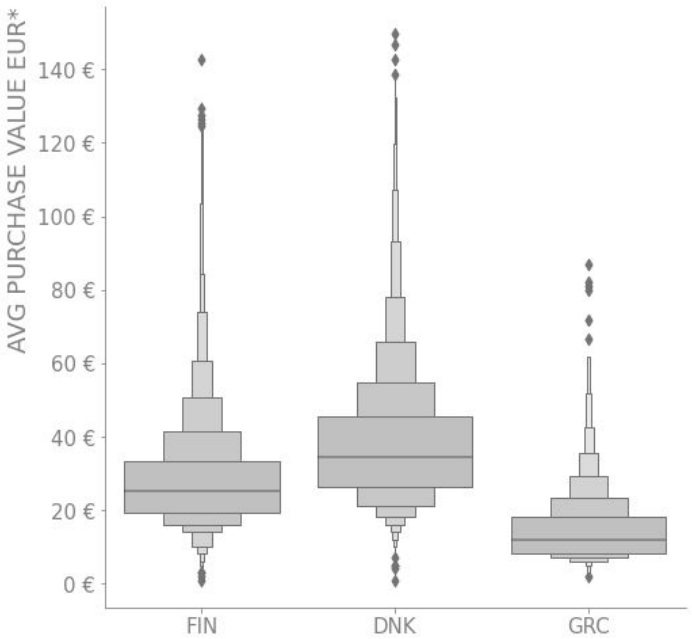
Total purchases value in EUR by registration country and groups for number of purchases



- For a business like Wolt the volume of purchases is more important than the total value of the purchases as I assume there is a fix fee for each transaction rather than a percentage of the transaction value.

Denmark has the highest **median** for the average purchase value of **34€**, followed by Finland with **25€** and Greece with **12€**.

Distribution of average purchase value in EUR by registration country



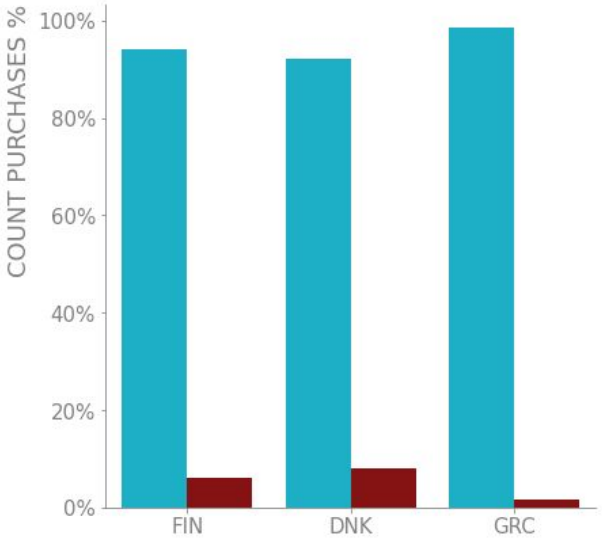
- Differences in the average purchase value between countries could be explained by different prices for food and groceries in the 3 countries rather than a different user behavior.
- Especially when comparing Greece with Denmark where the food is much cheaper.

*showing data only for AVG_PURCHASE_VALUE_EUR <= 150 to highlight differences between the 3 countries without the outliers

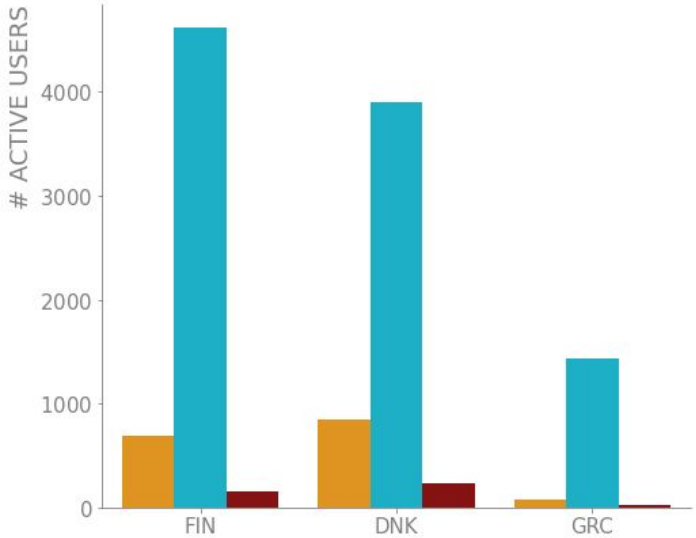
**see Appendix for explanation on how to read the boxen plots

Over **92%** of all purchases in all 3 countries were **deliveries**. With most users being **delivery only** users.

Percentage of **delivery** purchases vs. **takeaway** purchase by registration country

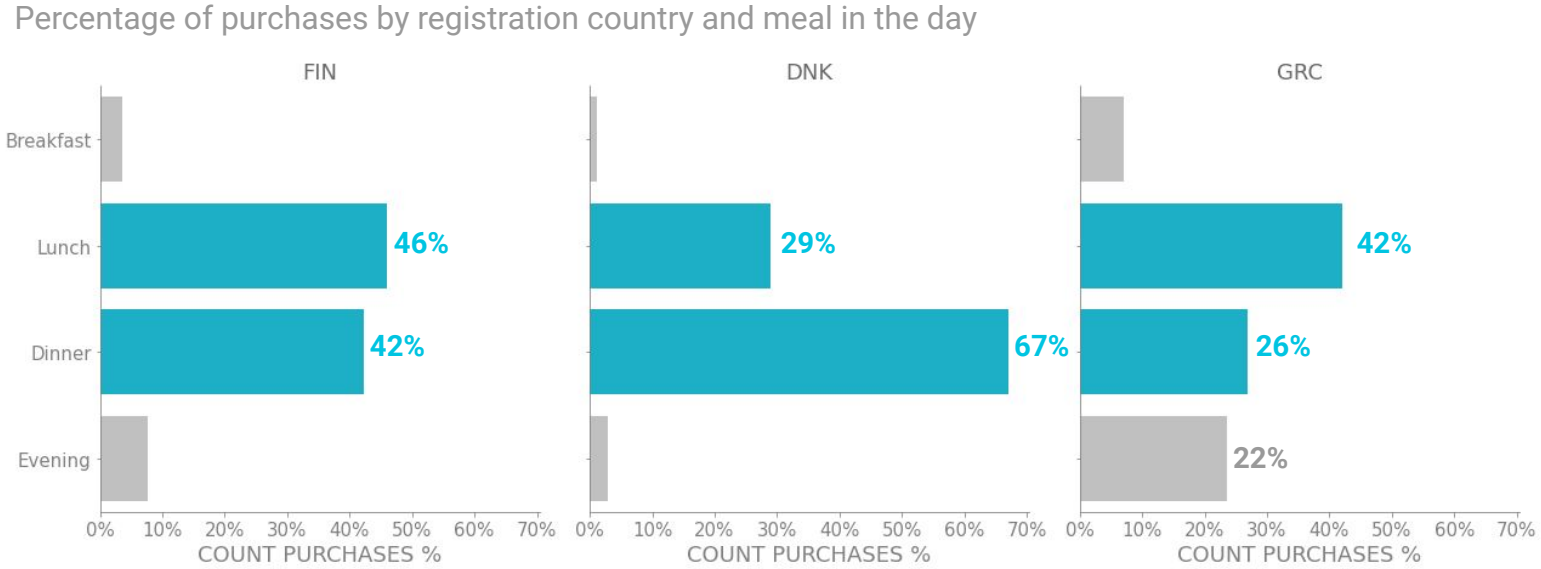


Active users with **both delivery and take away** purchases, **only delivery** purchase, **only takeaway** purchases



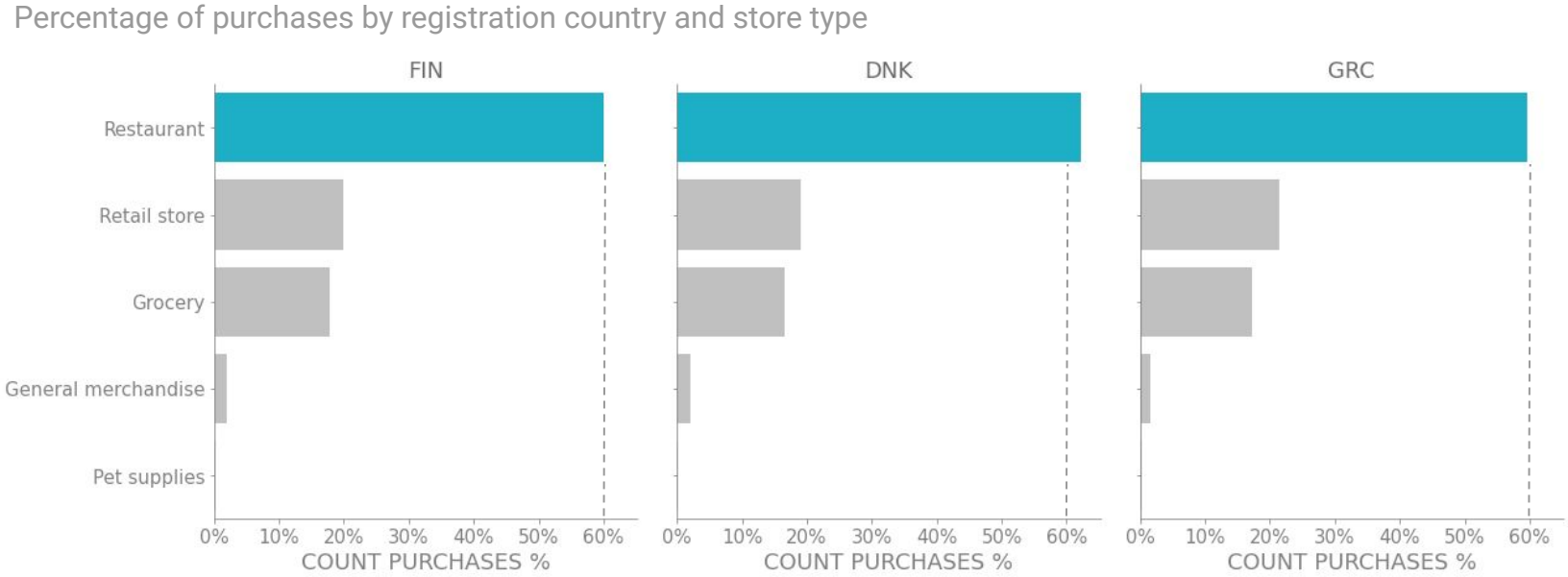
- It seems that if users use Wolt they will choose to deliver the order at home rather than picking it up themselves.

Purchases for **lunch** and **dinner** account for **88%** of all purchase in Finland, **96%** in Denmark and **68%** in Greece.



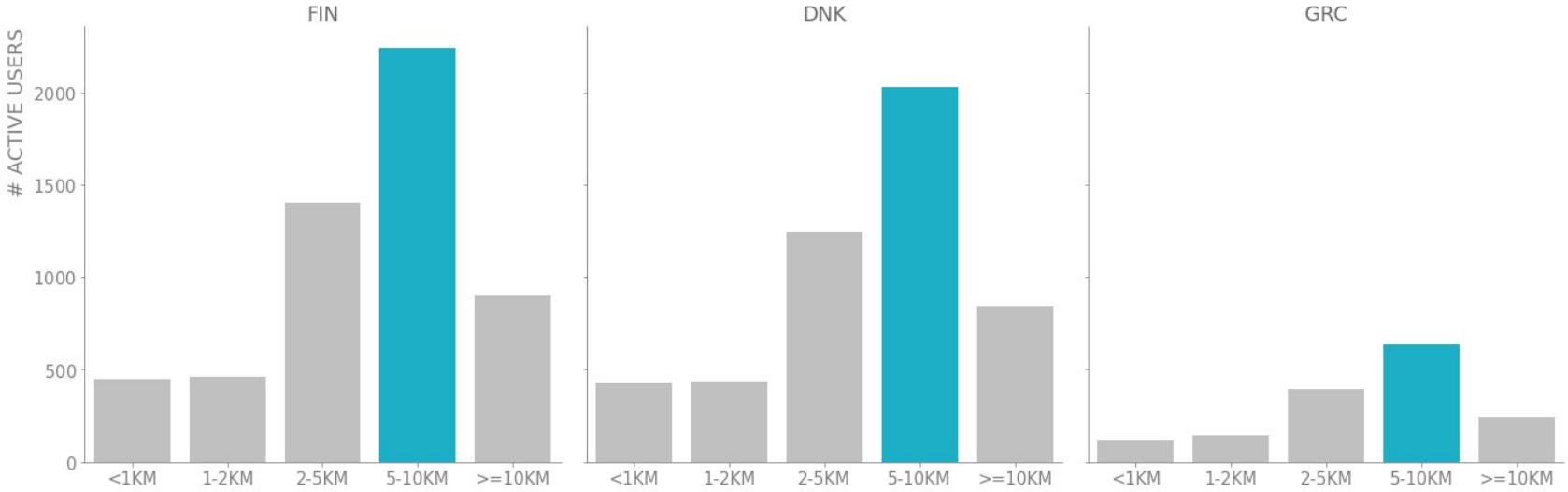
- Greece stands out with a significant share of purchases in the **evening**.
- In all countries there are few purchases for **breakfast**.

Purchases for **restaurant** account for **~60%** of all purchase in all 3 countries, followed by **retail store** and **grocery**.



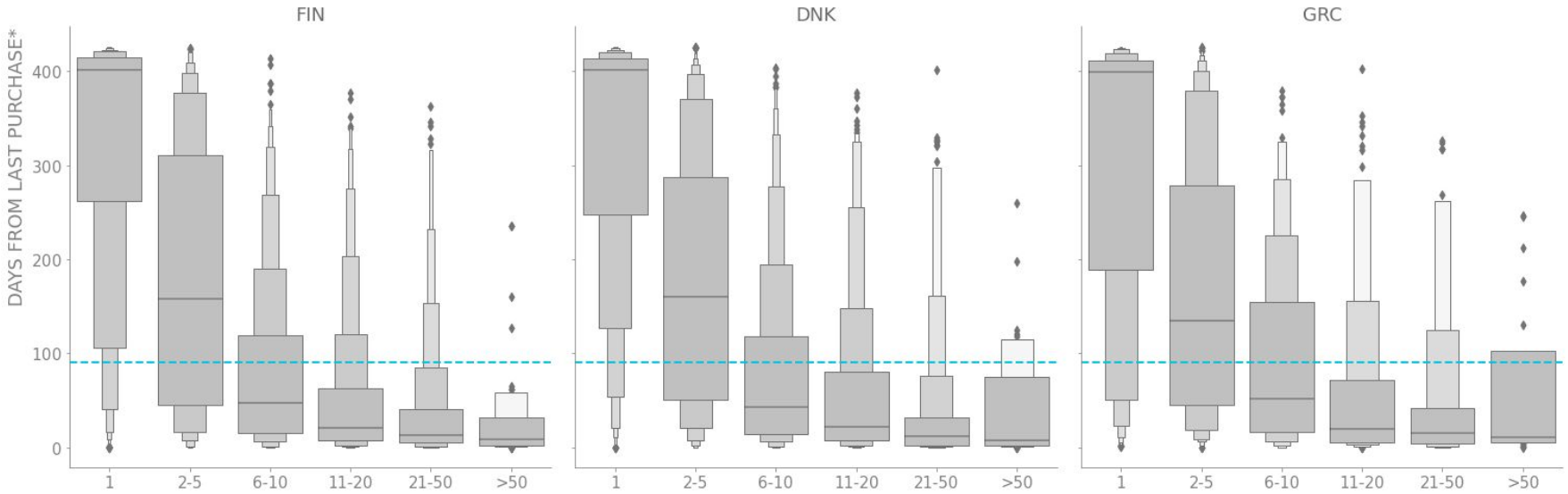
Most users order from venues within **10 KM** distance with a small percentage ordering from venues at more than 10 KM delivery distance.

Number of active users by registration country and average delivery distance in KM



Majority of users that made just 1 purchase within one year from registration have done so more than **90 days** ago while users with > 20 purchases have made at least 1 purchase within last **90 days**.

Distribution of days from last purchase by registration country and groups for number of purchases



*see Appendix for more details for logic behind the calculated column days from last purchase

User segmentation

From the provided data **purchase count** and **days from last purchase** columns were used for user segmentation.

- As Wolt is a very localized business the user segmentation is provided by **country** in the following 3 slides. Also I assume that the marketing team has local campaigns to engage and reactivate different kinds of users.
- The logic behind using these 2 columns is that they provide information on how active the users have been and when is the last time they engaged with Wolt. Most other columns are highly correlated with 1 of these 2 columns.
- User segmentation provided 4 segments: **inactive**, **one-time shoppers**, **casual shoppers**, **frequent shoppers**.

User segmentation - **Finland**

47%

INACTIVE

Users that registered but made no purchases within one year from registration.

- **6%** of inactive users have a **valid payment method** setup.
- Most preferred device is **android**.

18%

ONE-TIME SHOPPERS

Users that made just one purchase within one year from registration.

- **11%** of one-time shoppers made a purchase in the **last 90 days**.
- Most preferred device is **android**.
- Preferred store type **restaurant** only.
- One-time shoppers prefer home **deliveries** for **lunch** or **dinner** with a median average purchase value of **24€**.

31%

CASUAL SHOPPERS

Users that made between 2 and 20 purchases within one year from registration.

- **52%** of casual shoppers made a purchase in the **last 90 days**.
- Most preferred device is **android**.
- Preferred store type **restaurant** followed by **retail store** & **grocery**.
- Casual shoppers prefer home **deliveries** for **lunch** or **dinner** with a median average purchase value of **26€**.

4%

FREQUENT SHOPPERS

Users that made over 20 purchases within one year from registration.

- **89%** of casual shoppers made a purchase in the **last 90 days**.
- Most preferred device is **android**.
- Preferred store type **restaurant** followed by **retail store** & **grocery**.
- Frequent shoppers prefer home **deliveries** for **lunch** or **dinner** with a median average purchase value of **22€**.

User segmentation - **Denmark**

38%

INACTIVE

Users that registered but made no purchases within one year from registration.

- **4%** of inactive users have a **valid payment method** setup.
- Most preferred device is **ios**.

21%

ONE-TIME SHOPPERS

Users that made just one purchase within one year from registration.

- **10%** of one-time shoppers made a purchase in the **last 90 days**.
- Most preferred device is **ios** followed by **web**.
- Preferred store type **restaurant** followed by **retail store** & **grocery**.
- One-time shoppers prefer home **deliveries** for **dinner** with a median average purchase value of **34€**.

37%

CASUAL SHOPPERS

Users that made between 2 and 20 purchases within one year from registration.

- **49%** of casual shoppers made a purchase in the **last 90 days**.
- Most preferred device is **ios** followed by **web**.
- Preferred store type **restaurant** followed by **retail store** & **grocery**.
- Casual shoppers prefer home **deliveries** for **dinner** with a median average purchase value of **35**.

4%

FREQUENT SHOPPERS

Users that made over 20 purchases within one year from registration.

- **89%** of casual shoppers made a purchase in the **last 90 days**.
- Most preferred device is **ios** followed by **web**.
- Preferred store type **restaurant** followed by **retail store** & **grocery**.
- Frequent shoppers prefer home **deliveries** for **dinner** with a median average purchase value of **32€**.

User segmentation - Greece

49%

INACTIVE

Users that registered but made no purchases within one year from registration.

- **3%** of inactive users have a **valid payment method** setup.
- Most preferred device is **android**.

19%

ONE-TIME SHOPPERS

Users that made just one purchase within one year from registration.

- **16%** of one-time shoppers made a purchase in the **last 90 days**.
- Most preferred device is **android**.
- Preferred store type **restaurant** only.
- One-time shoppers prefer home **deliveries** for **lunch**, **dinner** or **evening** with a median average purchase value of **10€**.

28%

CASUAL SHOPPERS

Users that made between 2 and 20 purchases within one year from registration.

- **48%** of casual shoppers made a purchase in the **last 90 days**.
- Most preferred device is **android**.
- Preferred store type **restaurant** followed by **retail store** & **grocery**.
- Casual shoppers prefer home **deliveries** for **lunch**, **dinner** or **evening** with a median average purchase value of **13€**.

4%

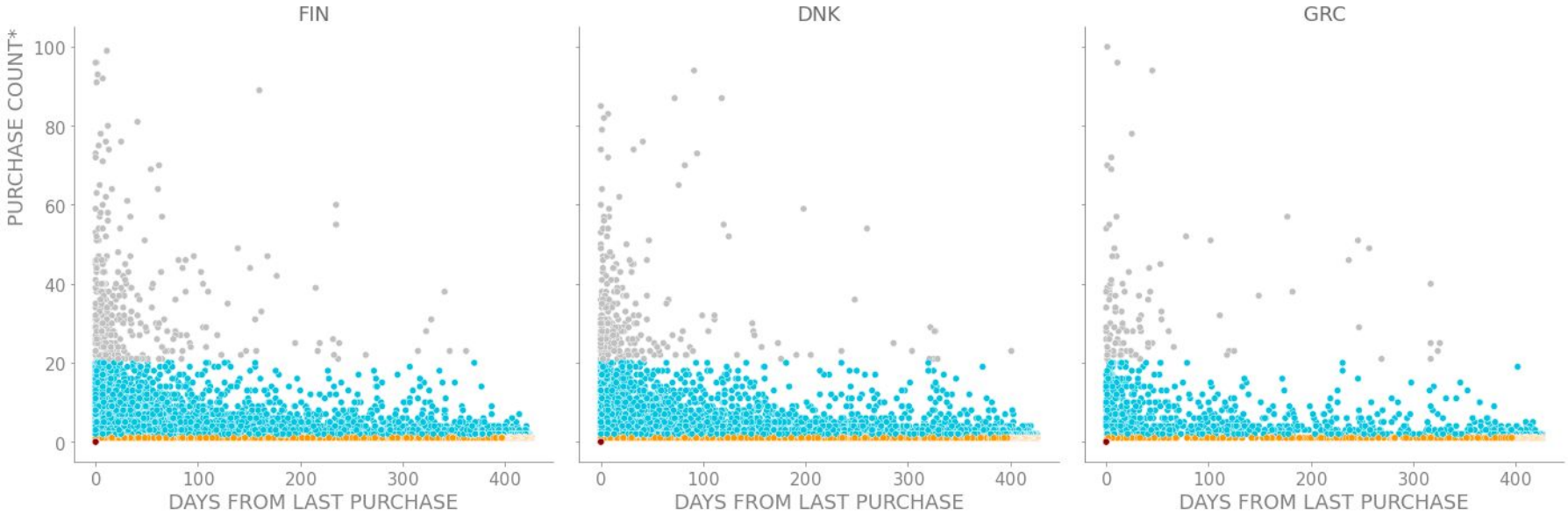
FREQUENT SHOPPERS

Users that made over 20 purchases within one year from registration.

- **81%** of casual shoppers made a purchase in the **last 90 days**.
- Most preferred device is **ios**.
- Preferred store type **restaurant** followed by **retail store** & **grocery**.
- Frequent shoppers prefer home **deliveries** for **lunch**, **dinner** or **evening** with a median average purchase value of **13€**.

Visualize the 4 different user segments

Purchase count vs. days from last purchase by registration country and user segment: **inactive**, **one-time shoppers**, **casual shoppers**, **frequent shoppers**.



*exclude outliers for purchase count > 100 for ease of visualization

Action: the marketing team should focus to reactivate **inactive** and **one-time shoppers**.

INACTIVE

- The goal is for inactive users to make their first purchase with Wolt.
- Prioritize users that have a valid payment method setup. However, this is quite a small share 3-6%.
- Majority of users prefer to use Wolt to order from restaurants, for lunch or dinner and have the order delivered at home.
- First purchase deals for restaurants and delivery, ideally for lunch or dinner should increase the chances for a purchase.

ONE-TIME SHOPPERS

- The goal is to reactivate one-time shoppers that made their purchase more than 90 days ago.
- For Finland & Greece we should target restaurant deals as this is what one-time shoppers are familiar with. In Denmark we could also target retail stores and groceries.
- Lunch and dinner deals for all countries with evening deals as well for Greece.
- Delivery deals would be appealing as well for one-time shoppers.

Comments

- Any attempt to use a clustering algorithm like k-means to segment the users failed due to very high correlation between columns. Even after trying to apply a dimensionality reduction method like Principle Component Analysis (PCA).
- For this type of analysis I would work closely together with the marketing team as they are experts when it comes to methods for engaging and re-activating different kinds of users.
- It would have been useful to have more data on the users themselves for example demographic data like age, income, gender, location etc. Right now the dataset only contains information about purchasing behaviour.
- This analysis could serve as a starting point until more data is available.
- Even though the user segmentation is not sophisticated together with the overall analysis can provide some guidance to the marketing team on which users to focus on and how to best engage them.

Appendix

Data wrangling

Registration country: As 97% of all registered users are from Finland, Denmark and Greece, the analysis focused only on these 3 countries. The logic behind this is that Wolt is a very localized business and we cannot draw any conclusions for rest of the countries where we have too little data.

Preferred device: I assume the preferred_device column represents the device where the users first registered. This makes sense for users that made no purchase but for users that actually made several purchases the preferred_device should be the device with most purchases. Also this column has 0.3% missing values which I assume come from data quality issues as we should know the device the user has used to register. To deal with these issues I created a new column preferred_device_corr where for users that made no purchase the data is the same and represents the device the user used to register, but for users that made at least one purchase this column represents the device with most purchases. Users with missing values for preferred_device_corr were excluded.

Preferred restaurant types: This column has 88% missing values for all users and 78% missing values when we select users that made at least one purchase so this column was not used in the analysis.

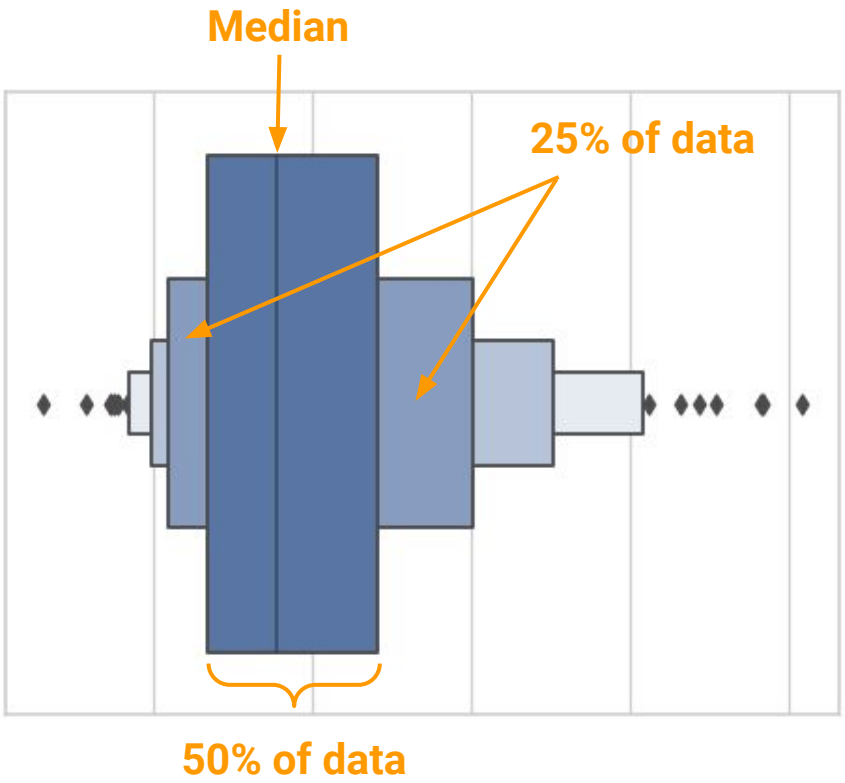
Late night purchases: Remove column as there were 0 purchases.

Data wrangling

Most common hour of the day to purchase: This column was ignored as the data did not make sense with an even number of purchases over the day. I assume 1.0 means 1am in the morning and 23.0 means 11pm. Also the data did not match the count of purchases by breakfast, lunch, dinner, evening.

Days from last purchase: I am assuming that this analysis would have been used some time after October 2020 to reactivate users that registered in September 2019. So this column was calculated as the difference in days between last purchase day and 2020-10-31.

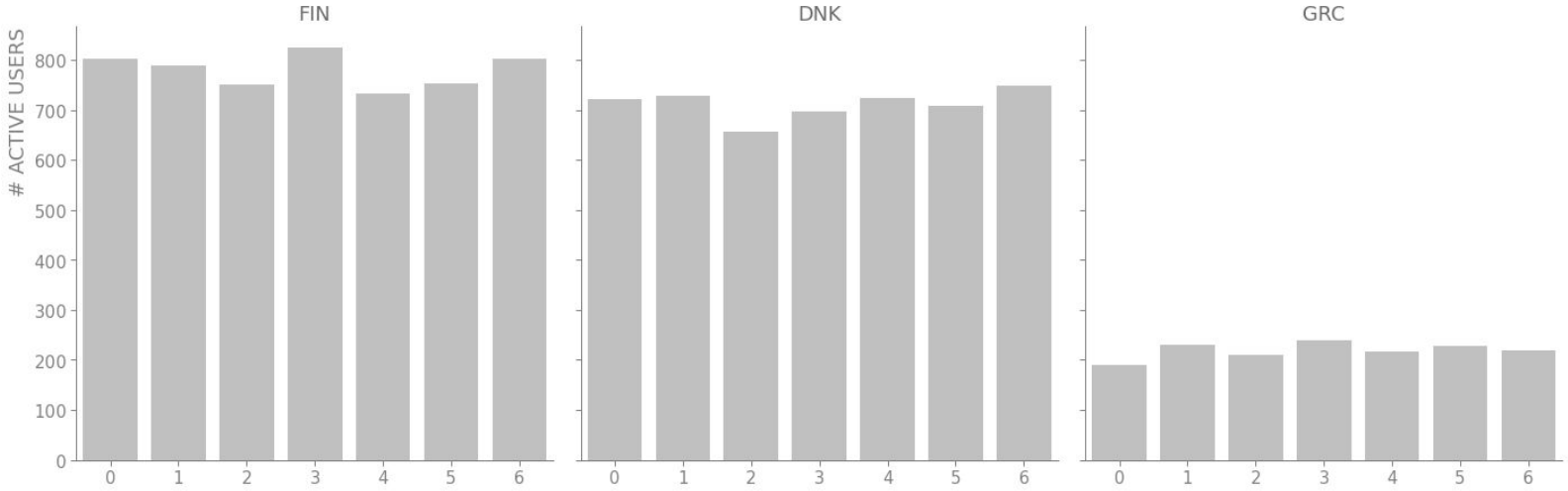
Understanding boxen plots



- Representation of a **distribution** in which all features correspond to actual observations.

There is no preferred weekday to purchase as the number of users is **evenly distributed** amongst most common weekday to purchase.

Number of active users by registration country and most common weekday to purchase



- This column does not contain much signal when it comes to user segmentation so it was ignored.