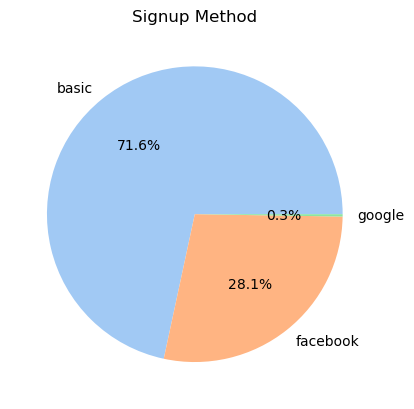
# About Dataset

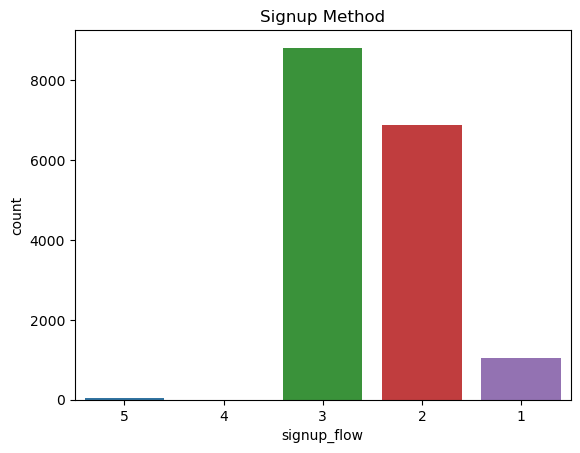
**Univariate analysis**

visualize the distribution of signup methods in the `users` dataset:

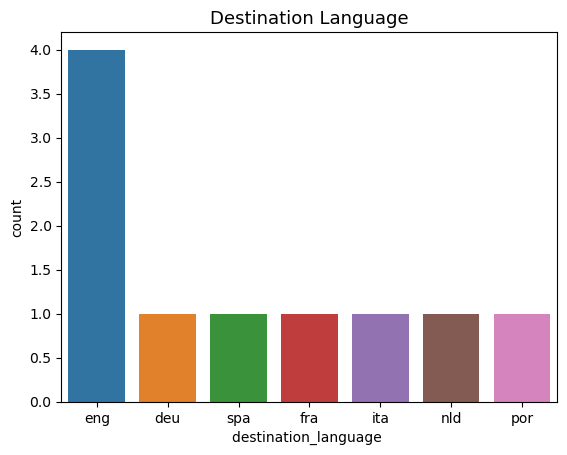


* signup\_flow: the page a user came to signup up from

View the distribution of user enrollment methods based on the signup\_flow column in the `users` dataset:

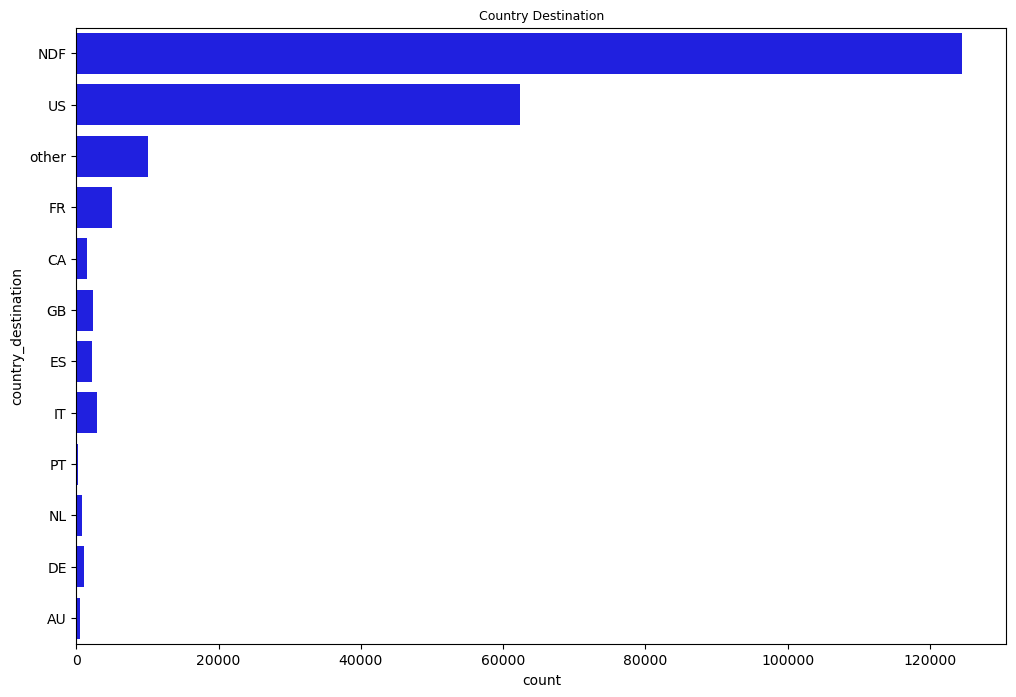


Visualize the destination language distribution based on the destination\_language column in the `users` dataset:



* Majority of users’ destination language is ‘eng’.

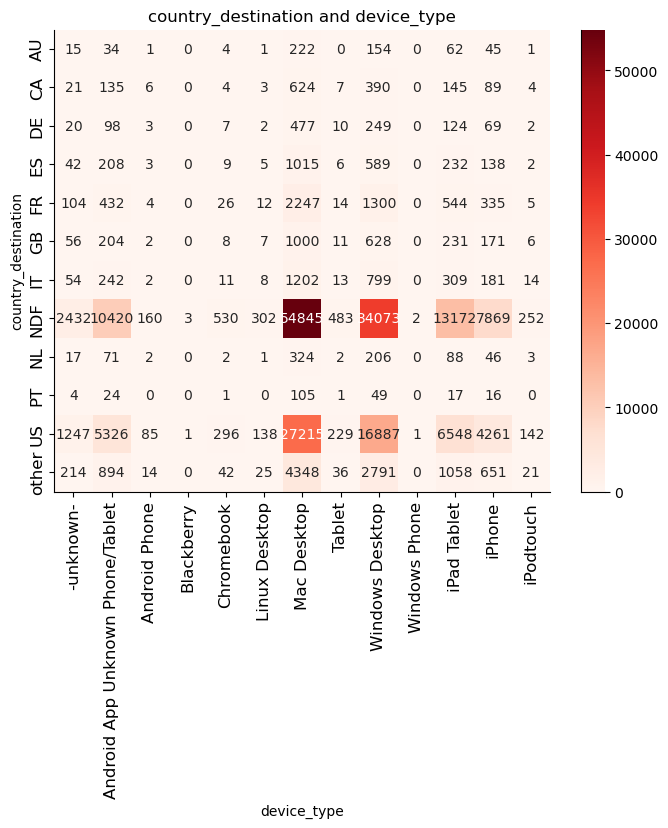
Visualize the country destination distribution in the `users` dataset:



* The Dataset is highly Imbalanced.
* Majority of users didn’t do any booking or travelled to the US only.

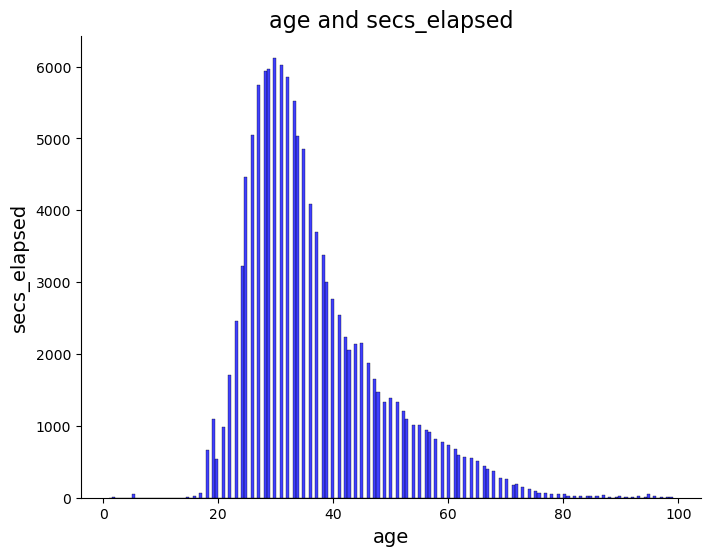
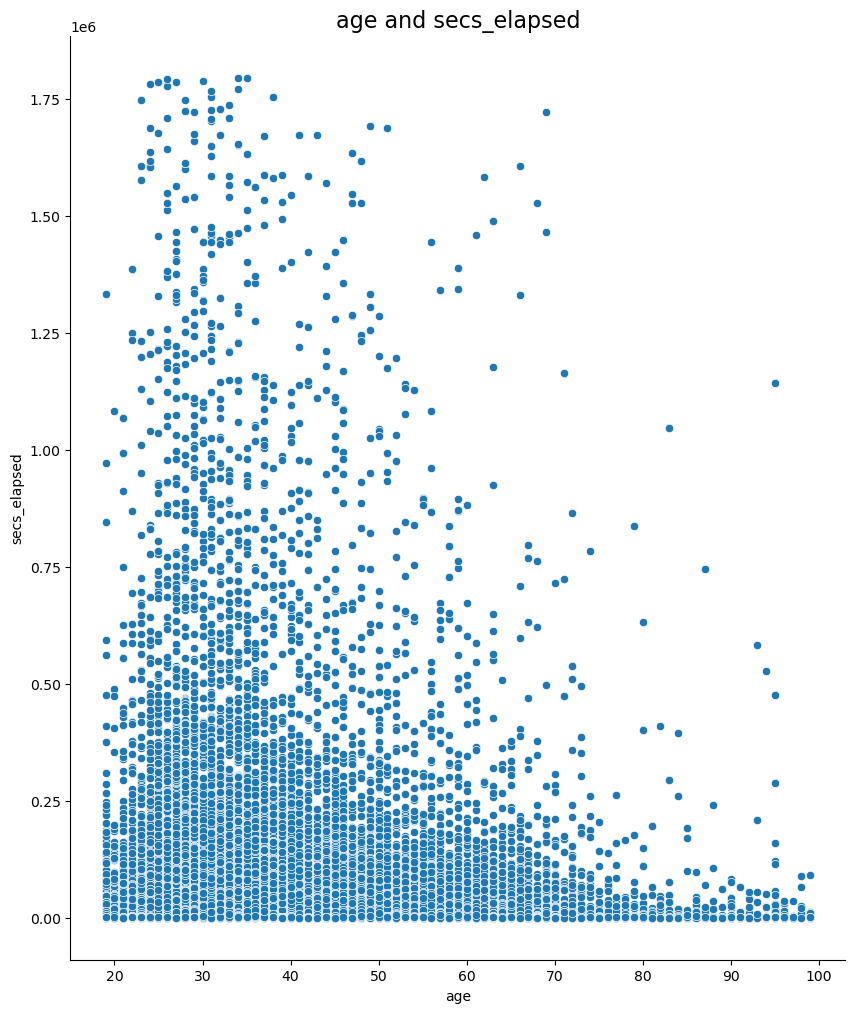
**Bivariate analysis**

Visualize the relationship between 'country\_destination' and 'device\_type' :



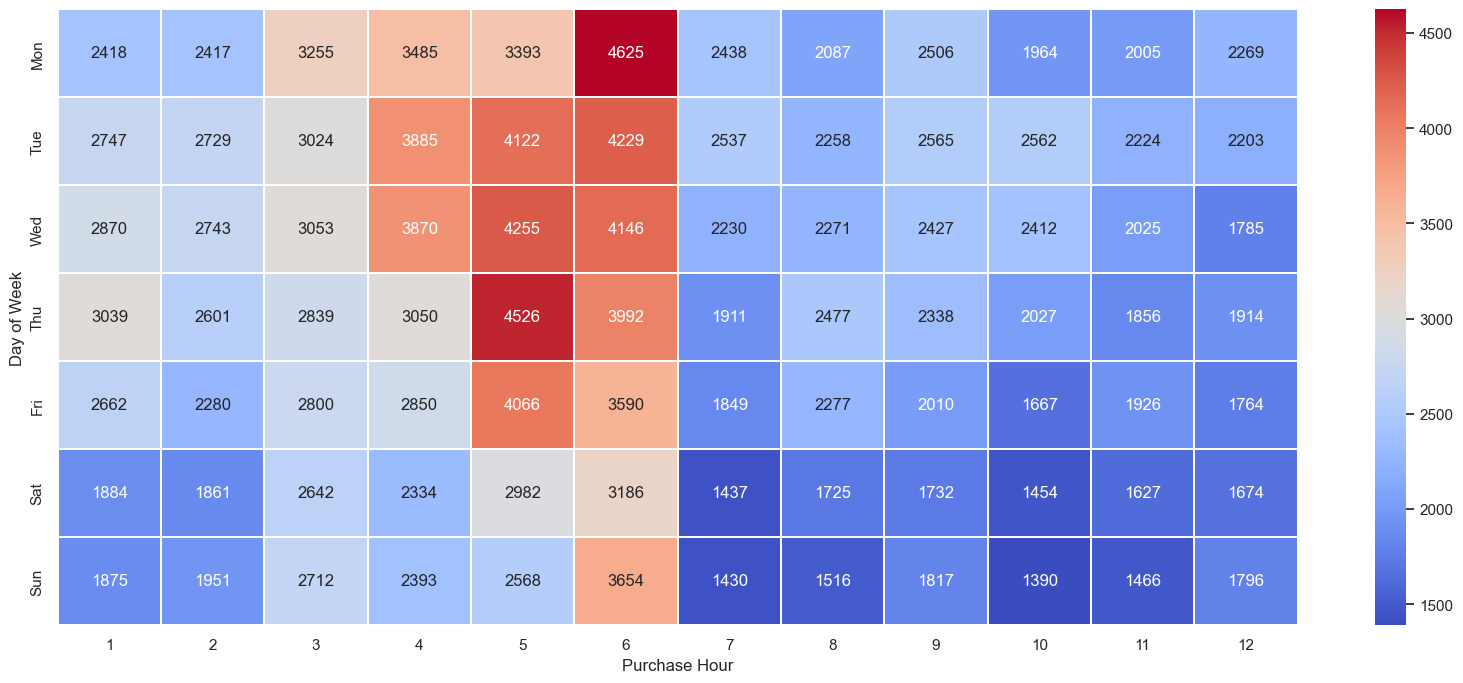
* The device type "Mac Desktop" is high in "NDF" country distination

Visualize the relationship between 'age' from the `users` dataset (filtered to include only ages below 100) and 'secs\_elapsed' from the `sessions` dataset:



* Ages between 30 and 40 are the majority of users' ages

visualize the count of account creations based on the day of the week and month:



* Months 5 and 6 are the most account creations

…

We can do other data analyses

…