

How to differentiate?



Mathieu Stremsdoerfer
July 11, 2018

1. Your profile and background information.

UP TO NOW

- ESCP EUROPE Master. MBA Insead in Marketing.
- International experience in digital, analytics, market research
- Serial entrepreneur and intrapreneur.
- Head of pricing and market research at Syngenta Lawn & Garden division

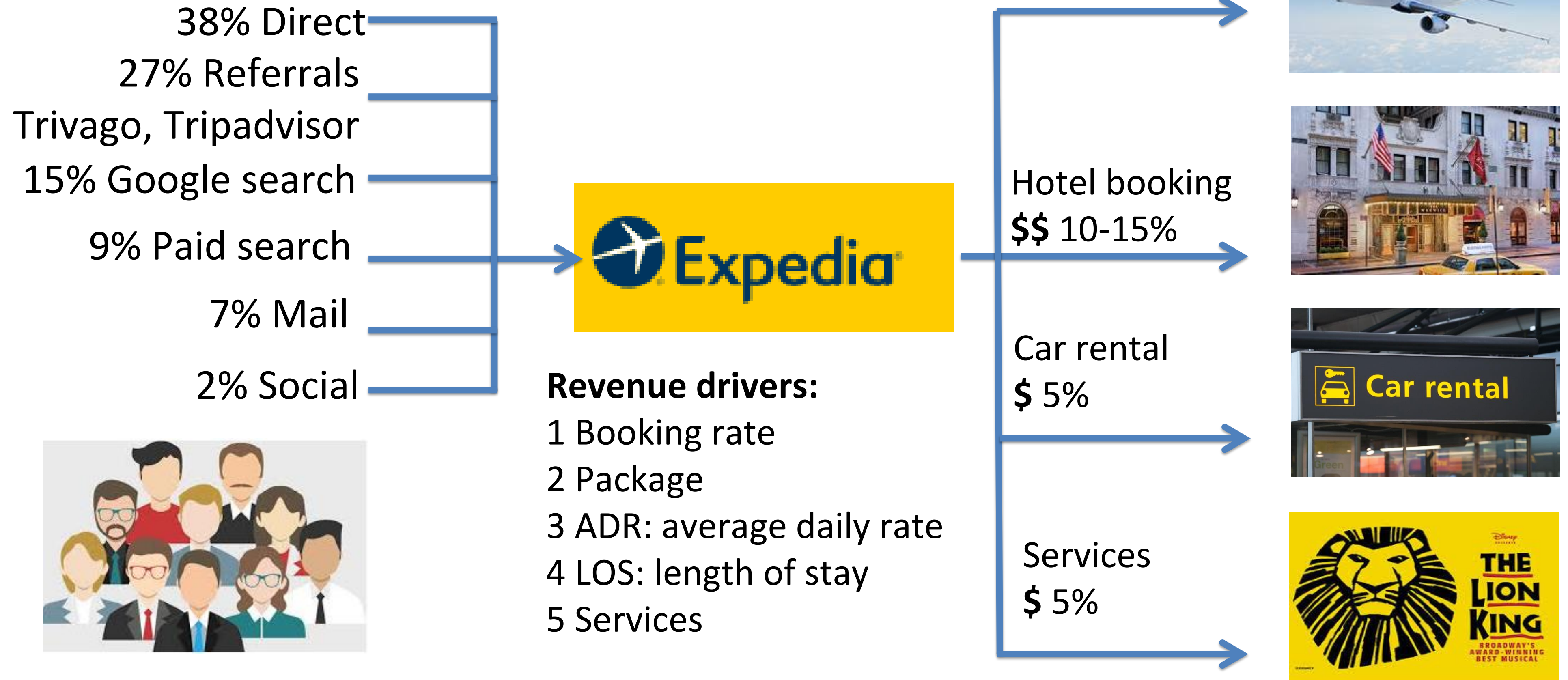
NOW:

- Student at Propulsion

FUTURE

- Connect data science with business

2. The business model



2. The challenge

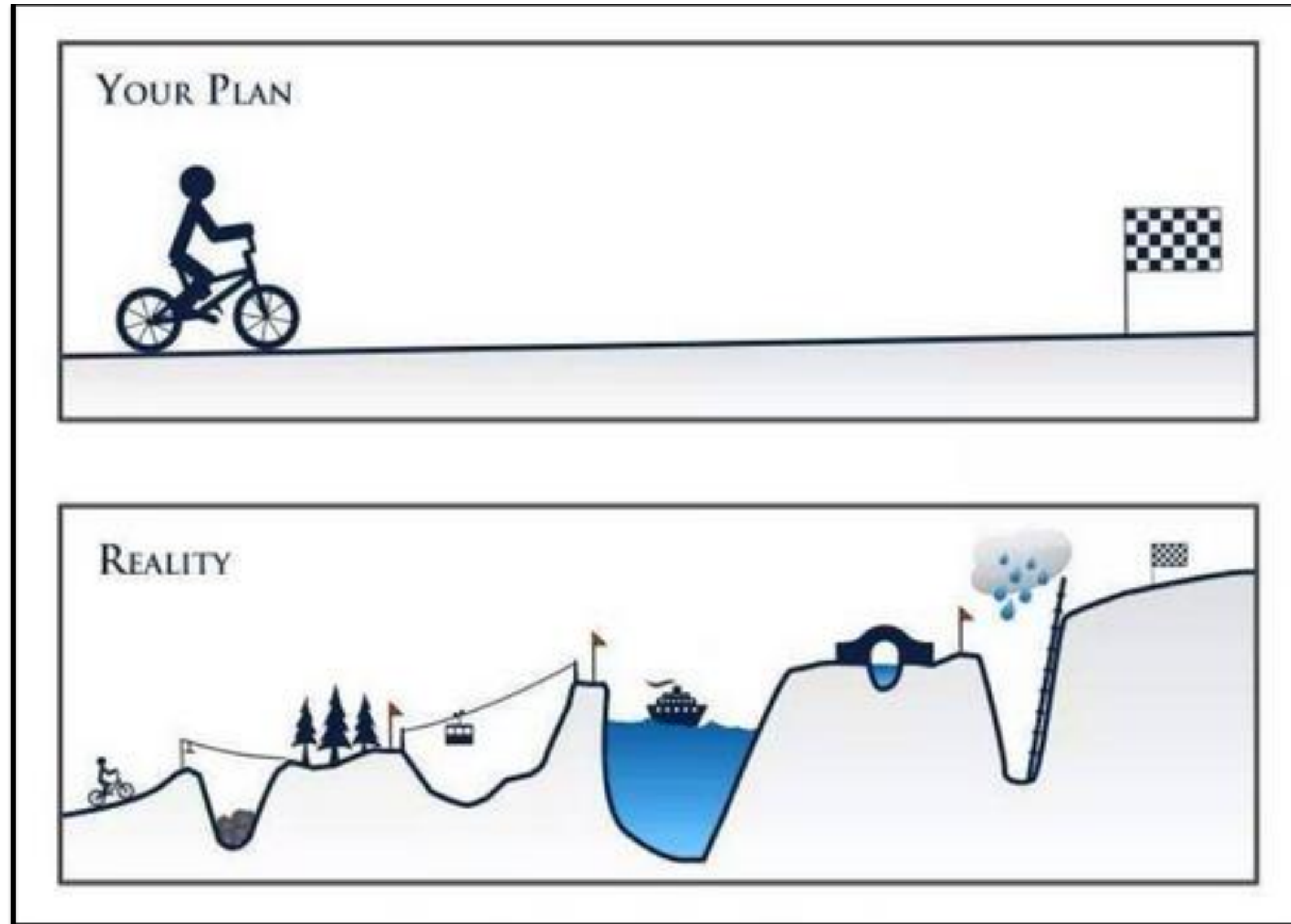
IT IS GETTING CROWDED OUT THERE...



GOAL:

“Use data to better understand customer needs, drivers of booking and create booking predictions”

3. Approach

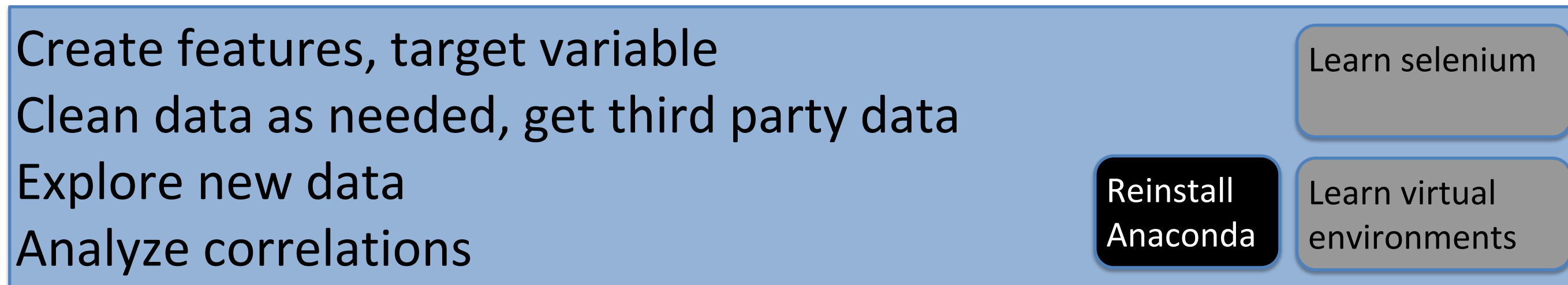


3. Approach

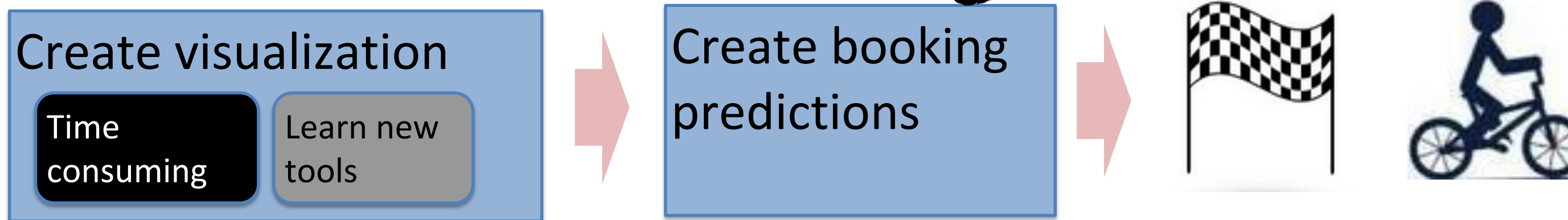
1st week



2nd week



3rd week



4. Execution: build the features from the sites entries

✈️

Flights

🏨

Hotels

✈️ 🏨

Bundle and Save

🚗

Cars

🚢

Cruises

🎫

Things to Do

🏠

Vacation Rentals

🔍

Discover

📍

Going to
New York, New York

✕

📅

Check-in
07/14/2018

📅

Check-out
07/18/2018

👤

Travelers
2 Adults, 0 Children, 1 Room

✈️

Add a flight

🚗

Add a car

Search

Examples:

Length Of Stay (LOS) = Date check-out – date check-in

Booking window (BW) = Date check-in – date searched

Business: = search (1 adult – 0 children during work days and hours) & travel (Monday- Friday)

“Bleisure”: search (1 adult – 0 children during work days and hours) & travel (Monday-Sunday)

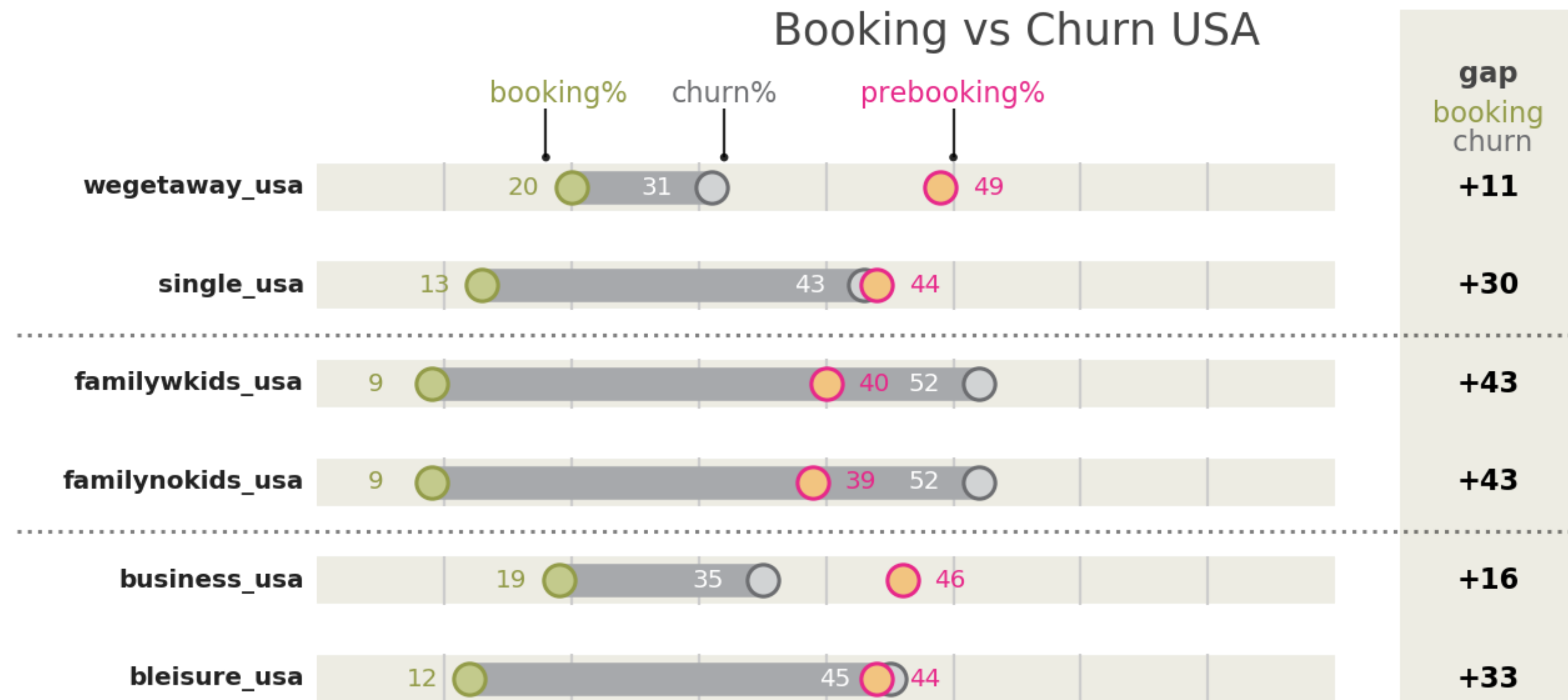
TOTAL: 35 new features in 9 categories

5. Definition: key binary variables

- **Booking:** for a specific user session,
Booking is 1 [hotel booked] or 0 [hotel not booked]
- **Churn:** for a specific user ID, destination (season)
Churn is 1 [sum of Booking = 0] else 0
- **Prebooking:** for a specific user session (season):
neither a churn nor a booking

$$\text{Prebooking} = 1 - \text{booking} - \text{churn}$$

5. Observations: Booking vs Churn vs Prebooking

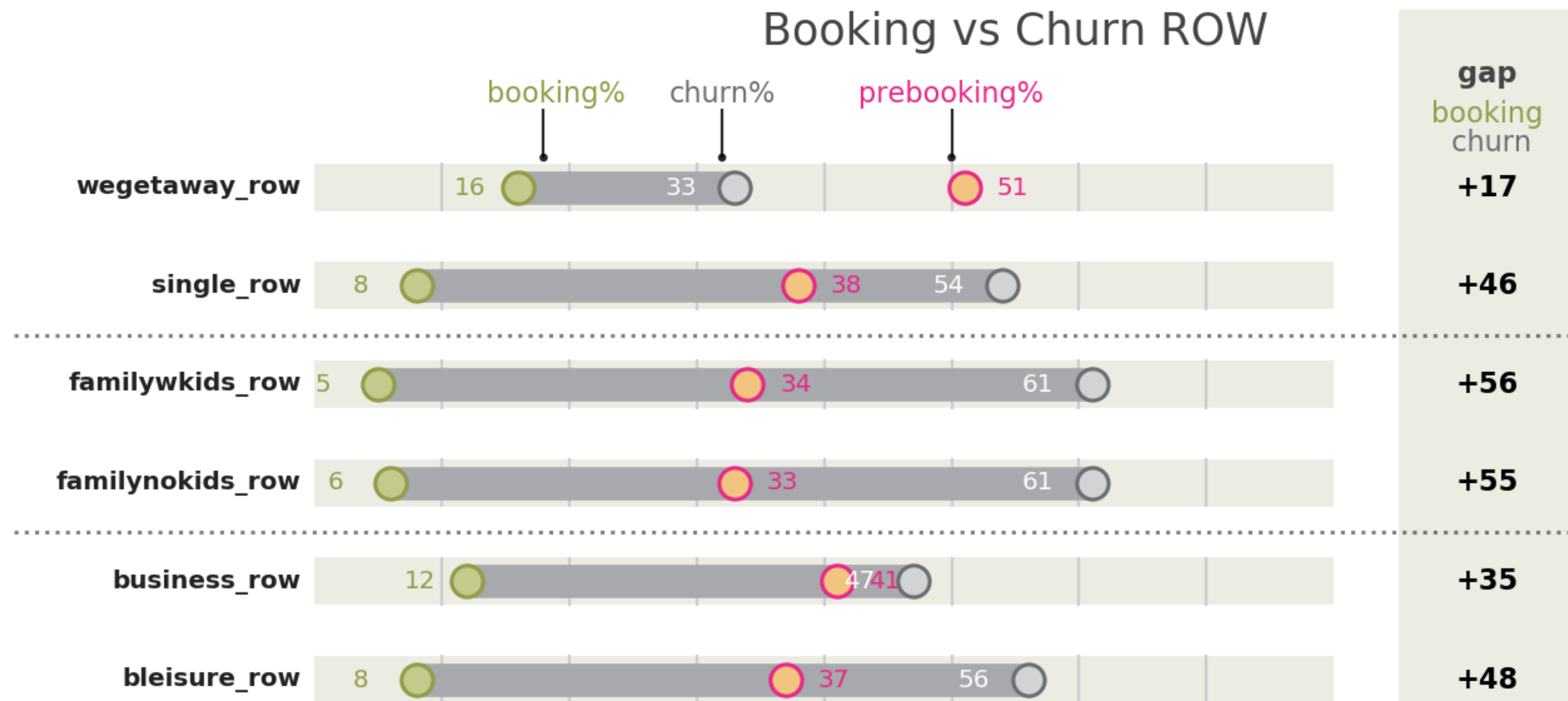


#1: User categories seem to “make sense”. Families underperform

#2: Booking positively correlated with prebooking

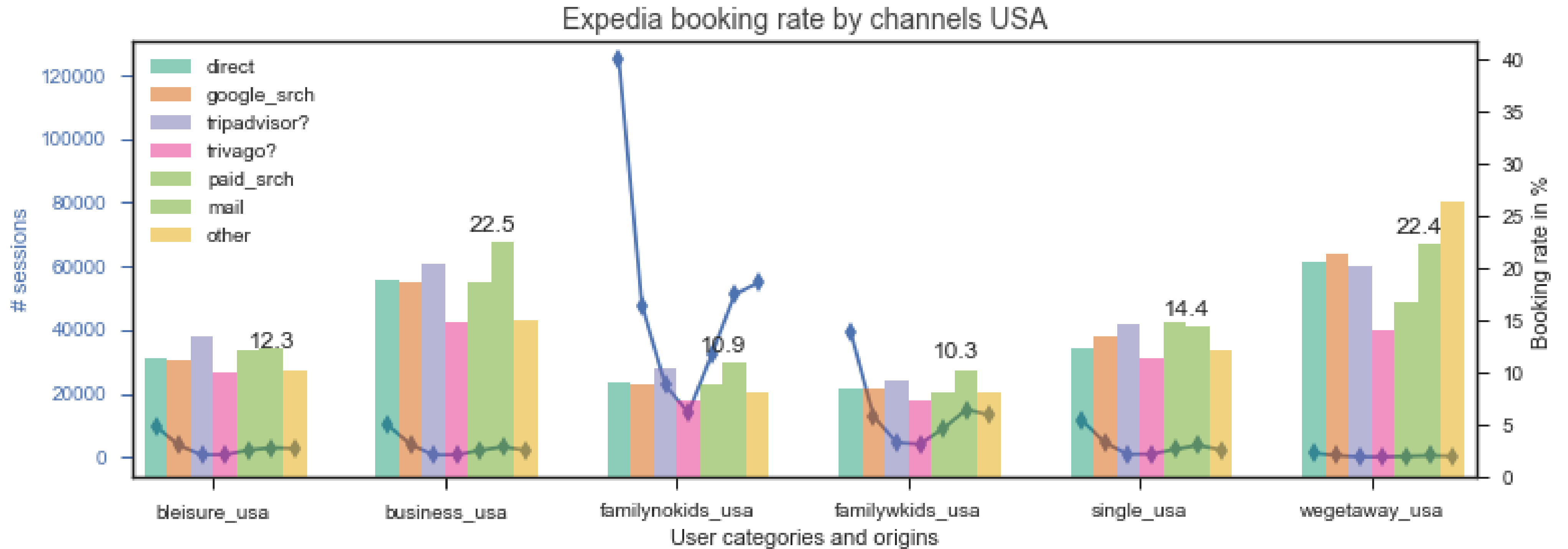


5. Booking vs churn



3: Same trends as USA, but lower booking rates overall

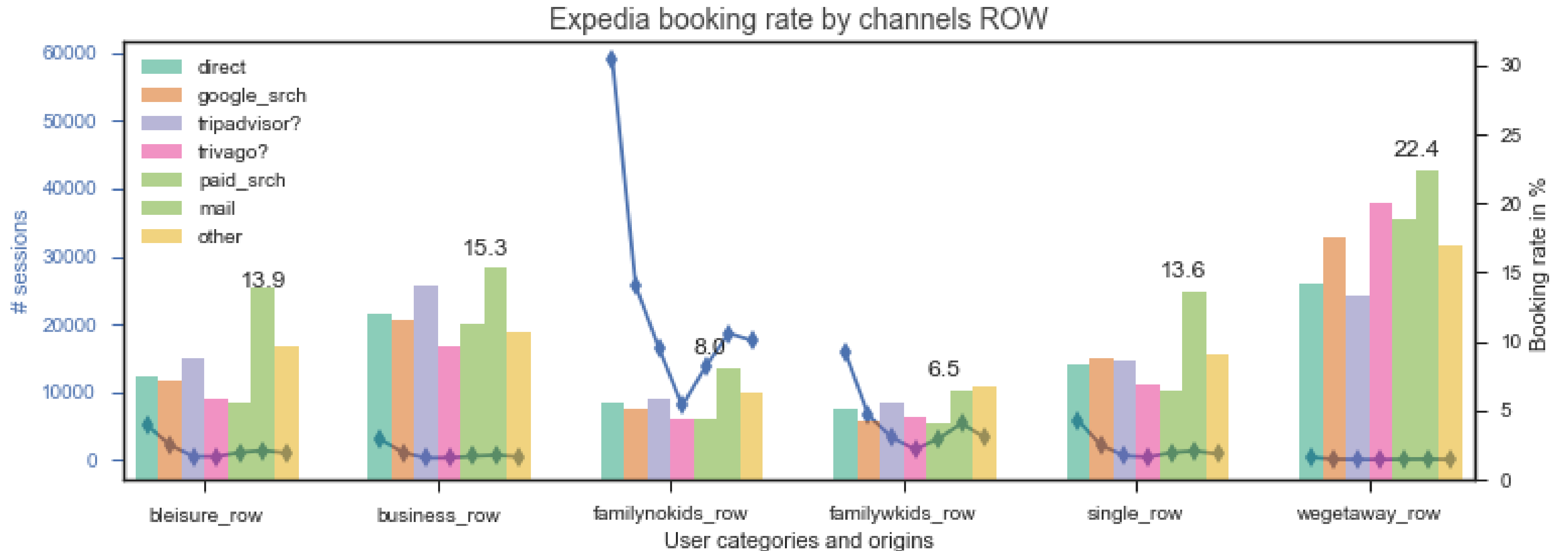
5. Booking vs channels



3: Booking rate varies by channel. GOOGLE around 15%

4: Most of the volume in family no kids

5. Booking vs channels



5 traffic ~ ½ of USA

6: same trends as USA

5. Income sensitivity versus packages

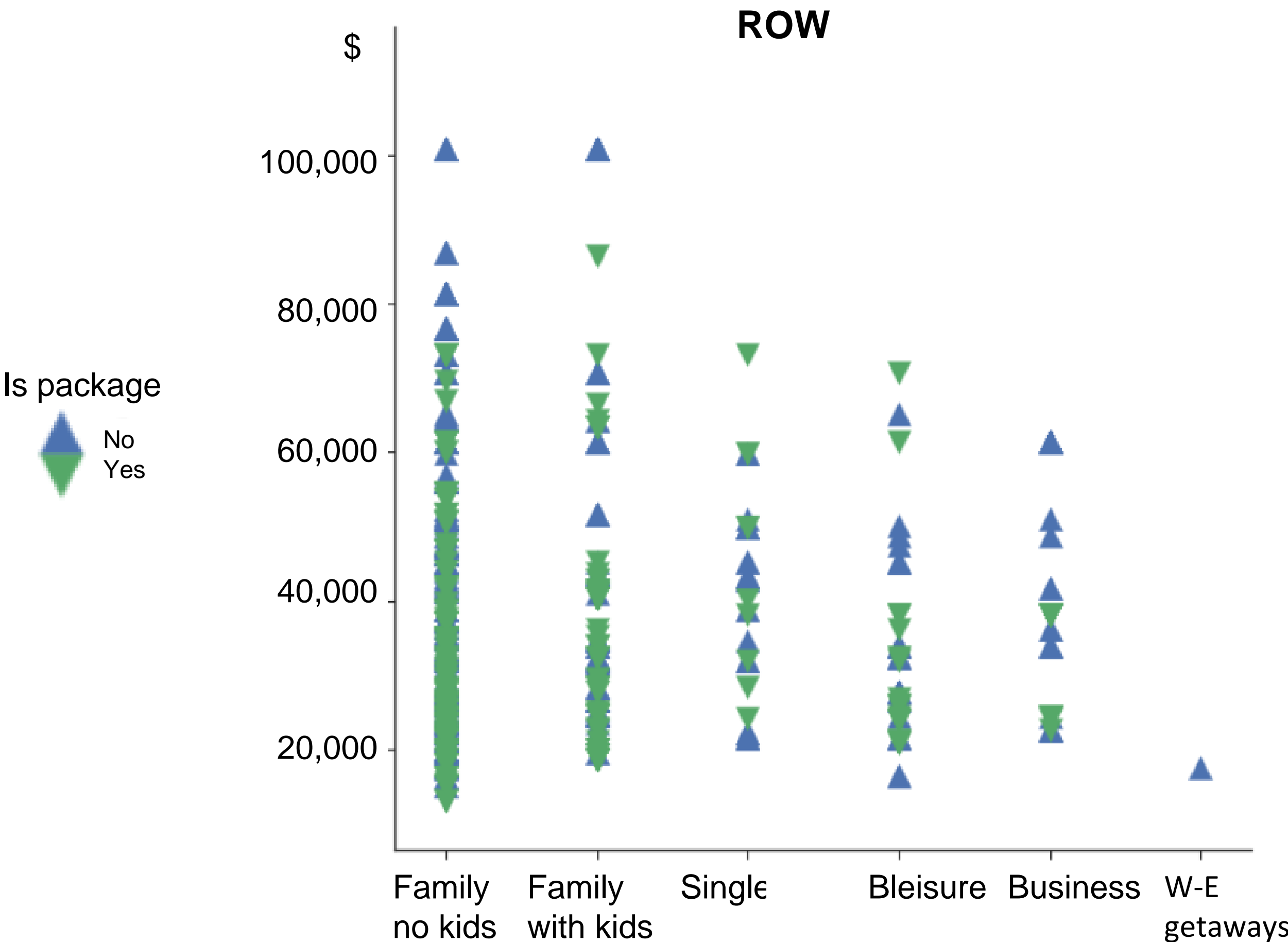


8: Clear opportunities for packages for budget travelers

Opportunity to geo-localize campaigns



5. Income sensitivity versus packages



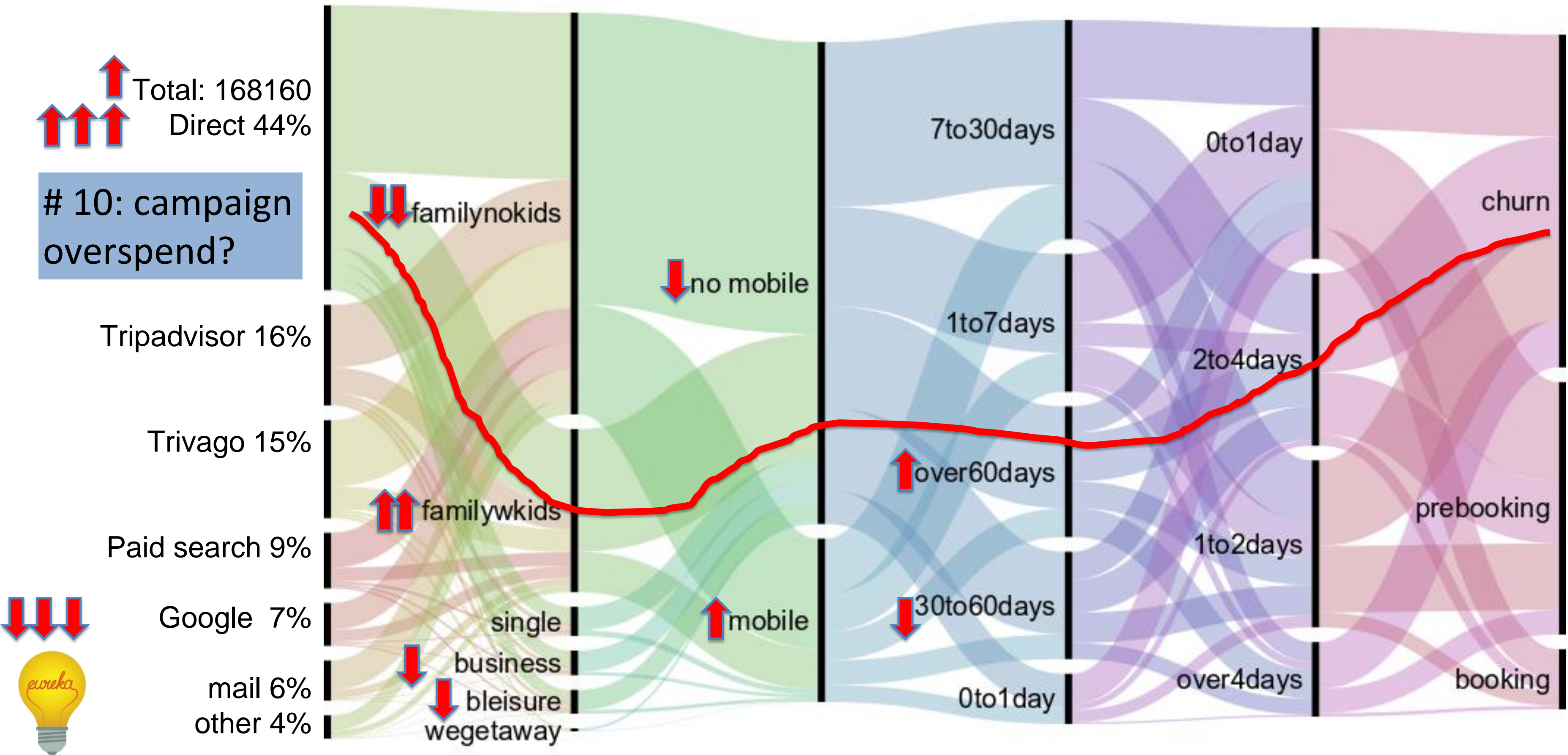
9: same!

Note:
Same correlation not found
with hotel stars and price
bands

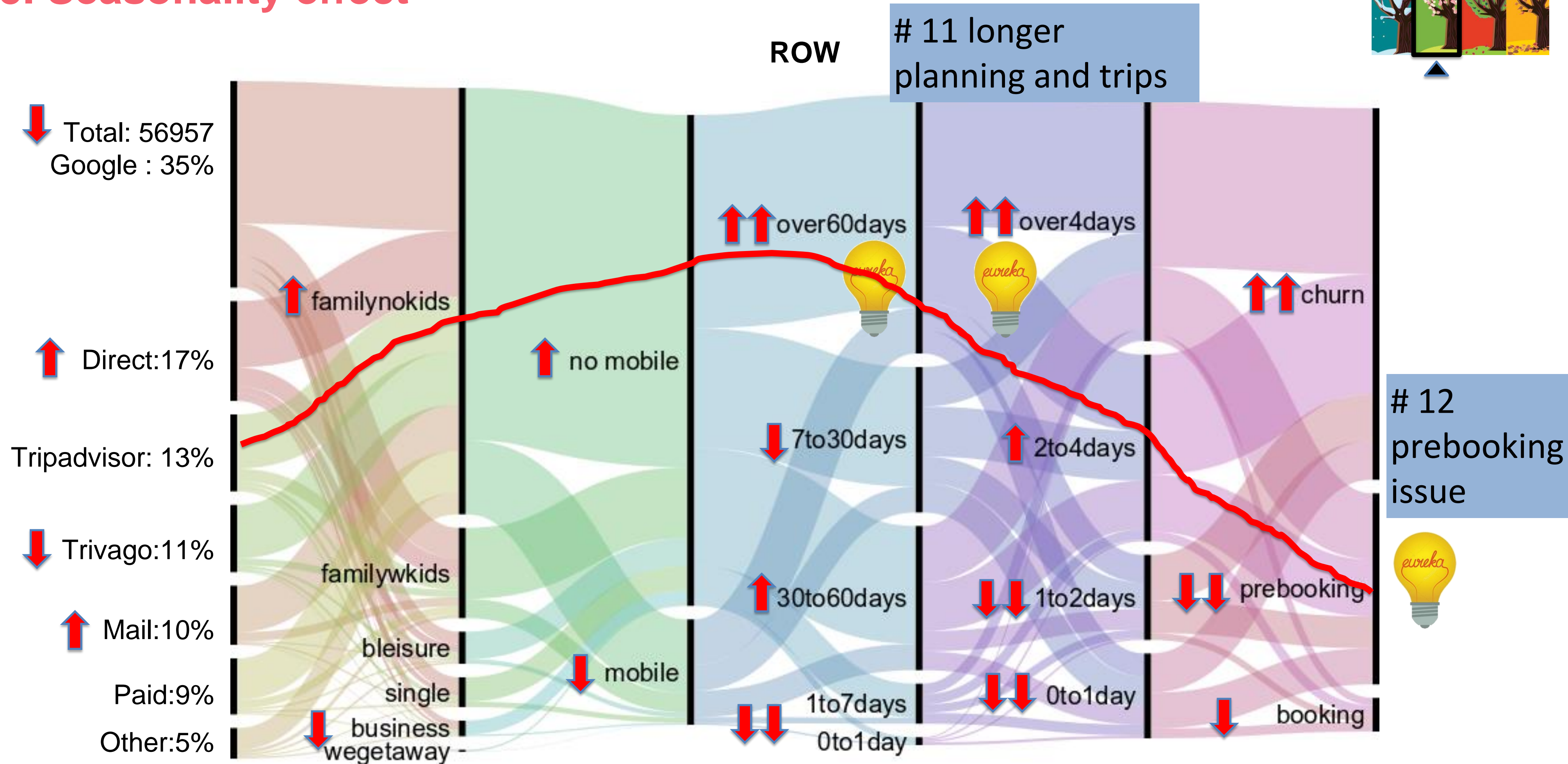
5. Seasonality effect



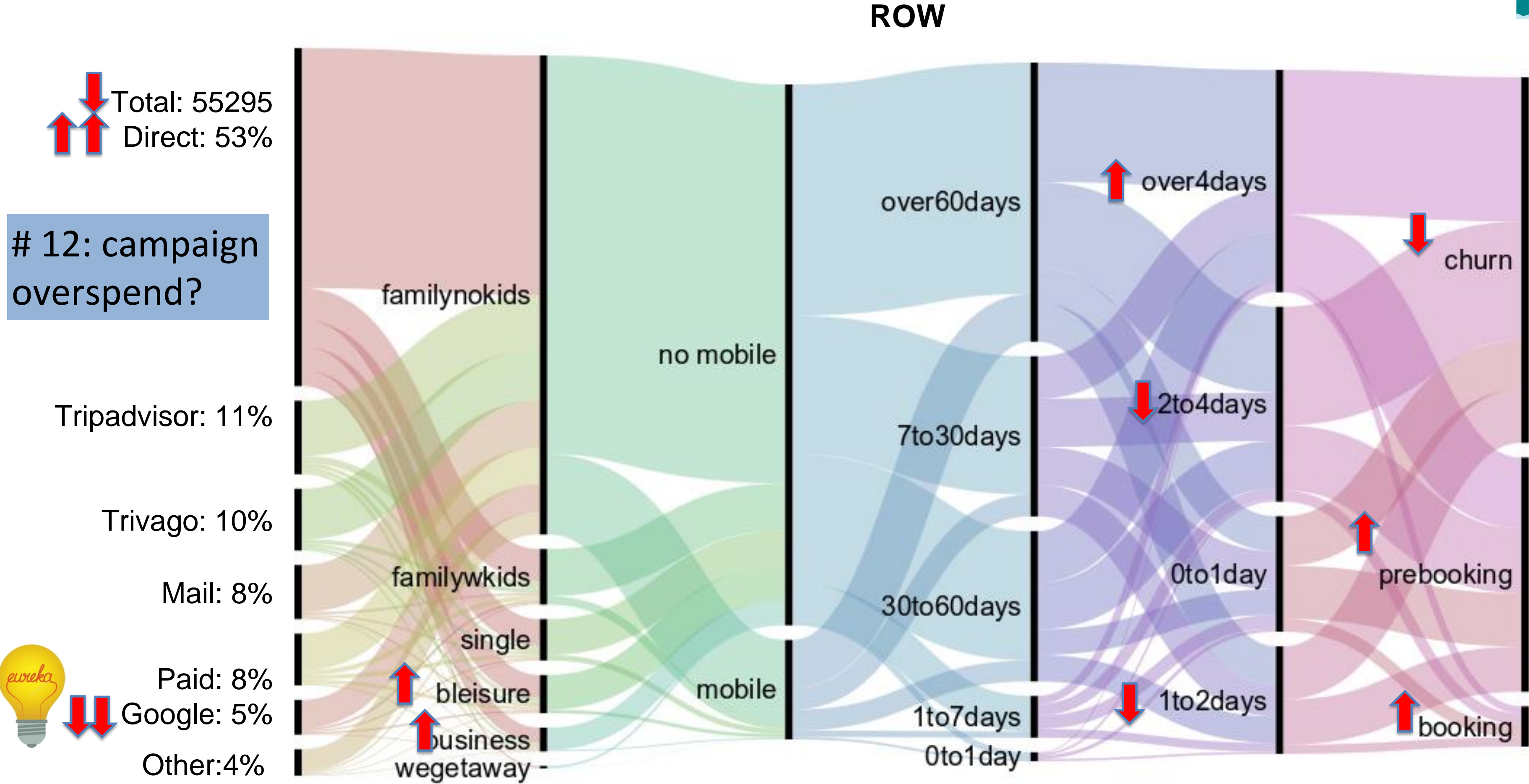
USA



5. Seasonality effect



5. Seasonality effect



5. Hotel booking predictions

Methodological Approach

Business problem

We want to predict if a user will **churn, prebook, or book in a given season**

Definition of the problem

A supervised –**multi class** classification problem with imbalanced classes and mixed variables (binary, numeric, categorical) with time series data

Preparation

- Divide dataset into **four seasons**
- Split train and test

Feature selection (kstratified folds, Man Whitney , collinearity, ridge regression)

Modelling: Logistic regression, KNN, Random Forest, Naïve Bayes, SVM, Gradient Boost,

Conclusions:



Quick win:

- Define user categories and session features per season to better understand needs
- Geo target campaigns based on IP/income – user categories
- Increase mobile offerings for last minute bookings



Tactical:

- Investigate Expedia seasonal campaign spent for possible optimization
- **Use booking/churn/prebooking prediction to identify “real churn” and offer help and attractive offers**



Strategic:

- Investigate root causes of low prebooking and BR for family and ROW
 - ex: is Expedia hotel offering adequate for families?
 - ex: should Expedia develop planning tools for ROW travelers?

Outlook

To go further:

- Complete multiclass classification
- Inquire user category stability over time and user loyalty
- Price data (quantitative) and multi year would be helpful for predictions

Perspective: marketing themes also applicable to other businesses:

- Fit for purpose segmentation
- Traffic drivers - channels
- Conversion efficiency: churn vs booking vs prebooking
- Price sensitivity - geo localization
- Seasonality
- Local vs international
- Campaign spend efficiency

Thank you!



{Propulsion}

EVEN THE
GREATEST WAS
ONCE A
BEGINNER. DON'T
BE AFRAID
TO TAKE THAT
FIRST STEP.



Mathieu Stremstoerfer



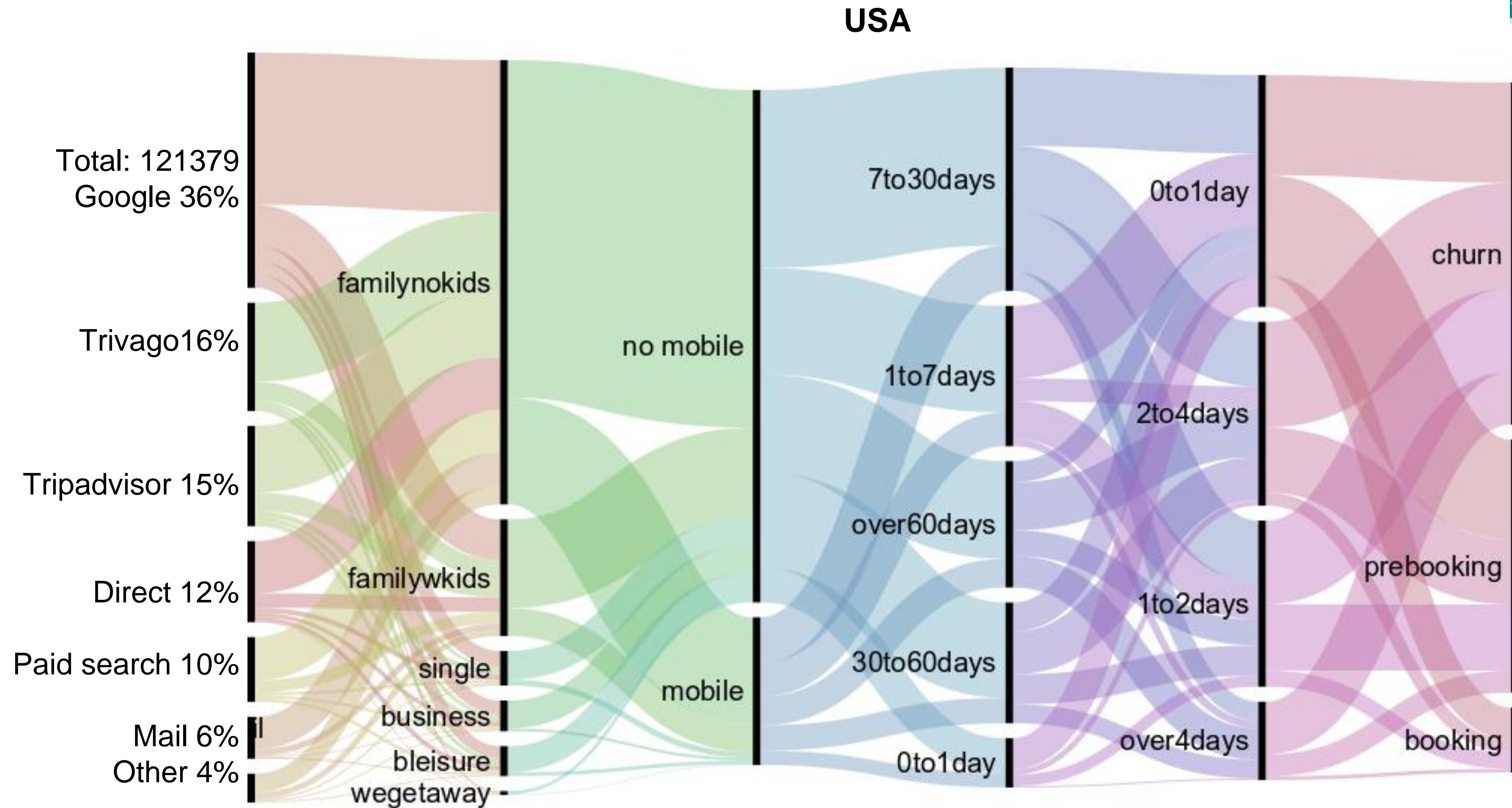
mstrems@gmail.com



www.linkedin.com/in/mstrems/

APPENDIX

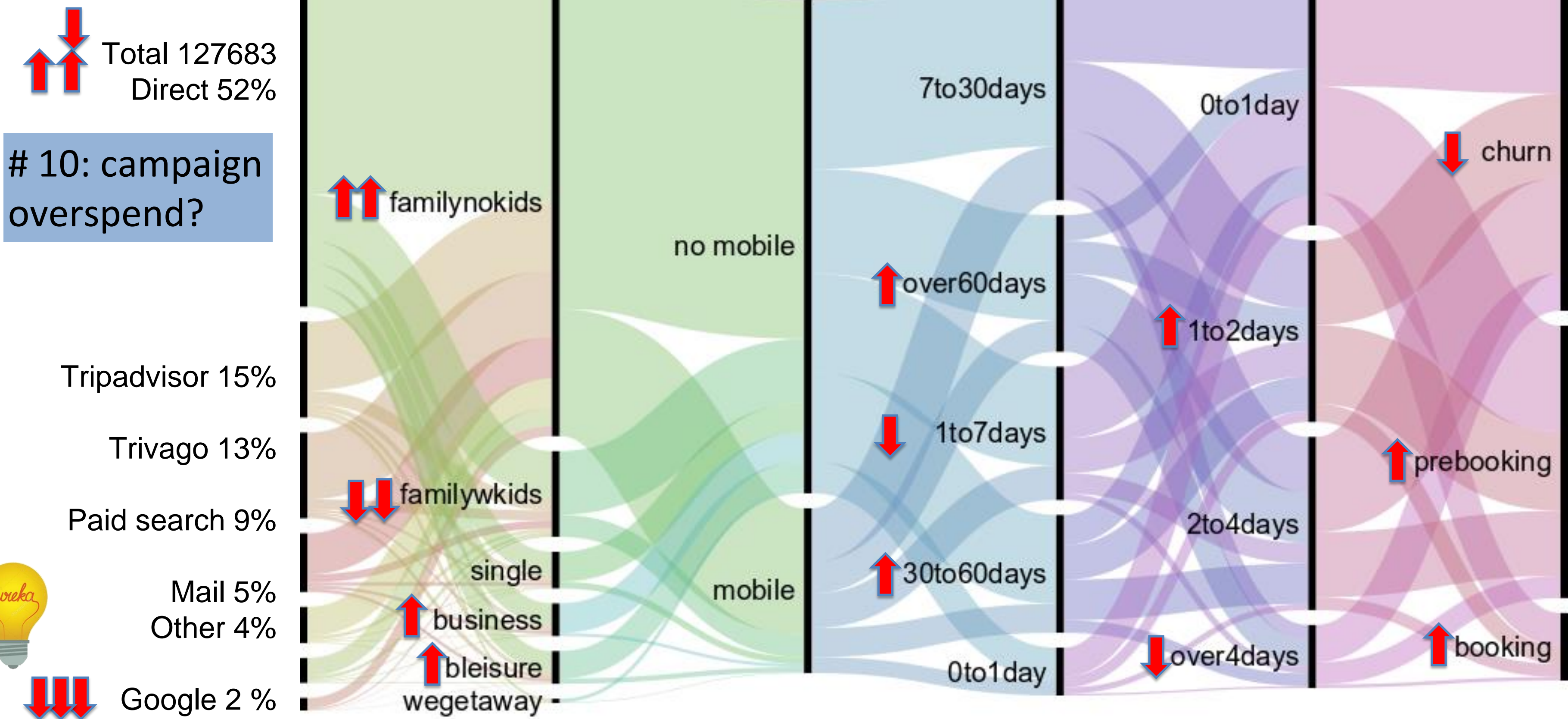
5. Seasonality effect



5. Seasonality effect



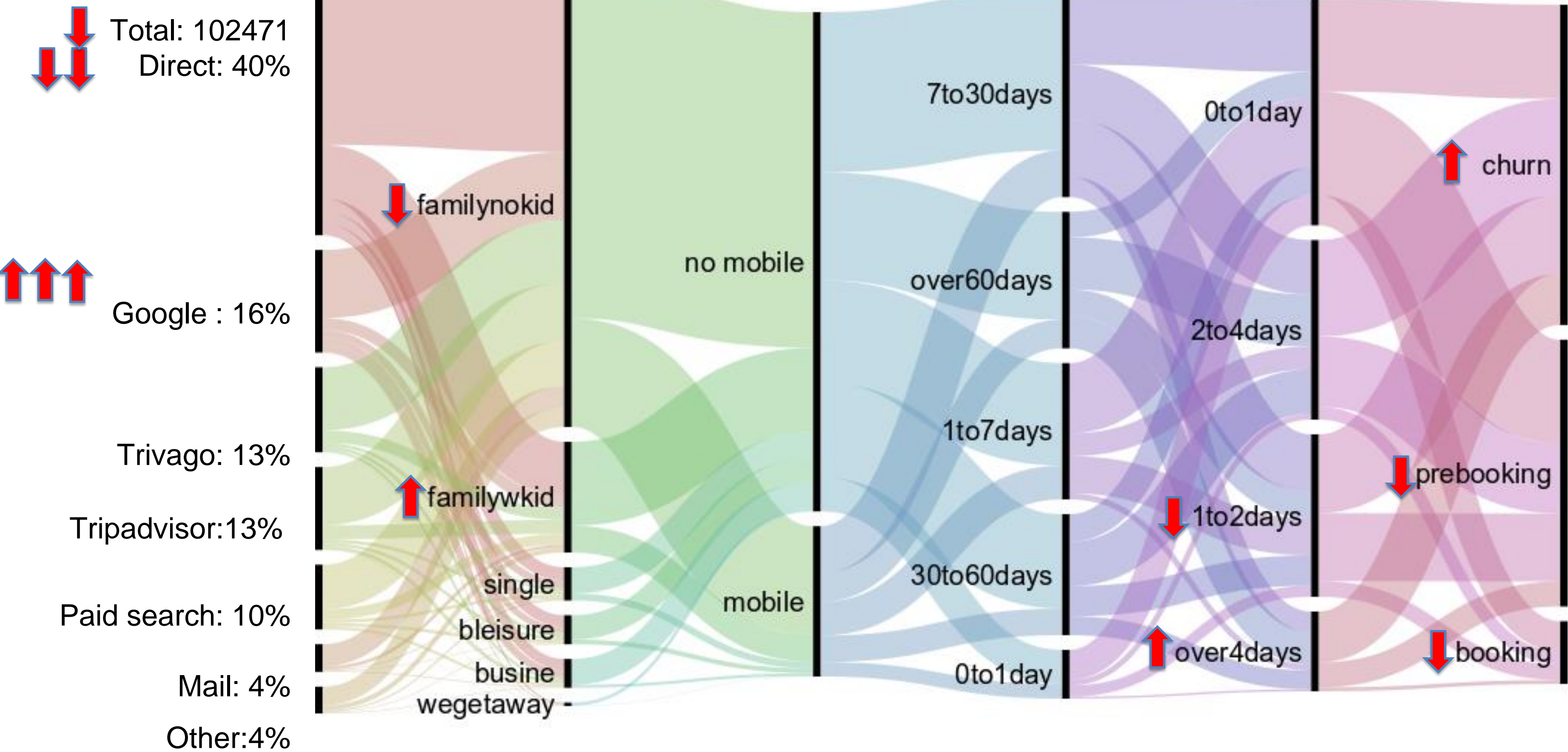
USA



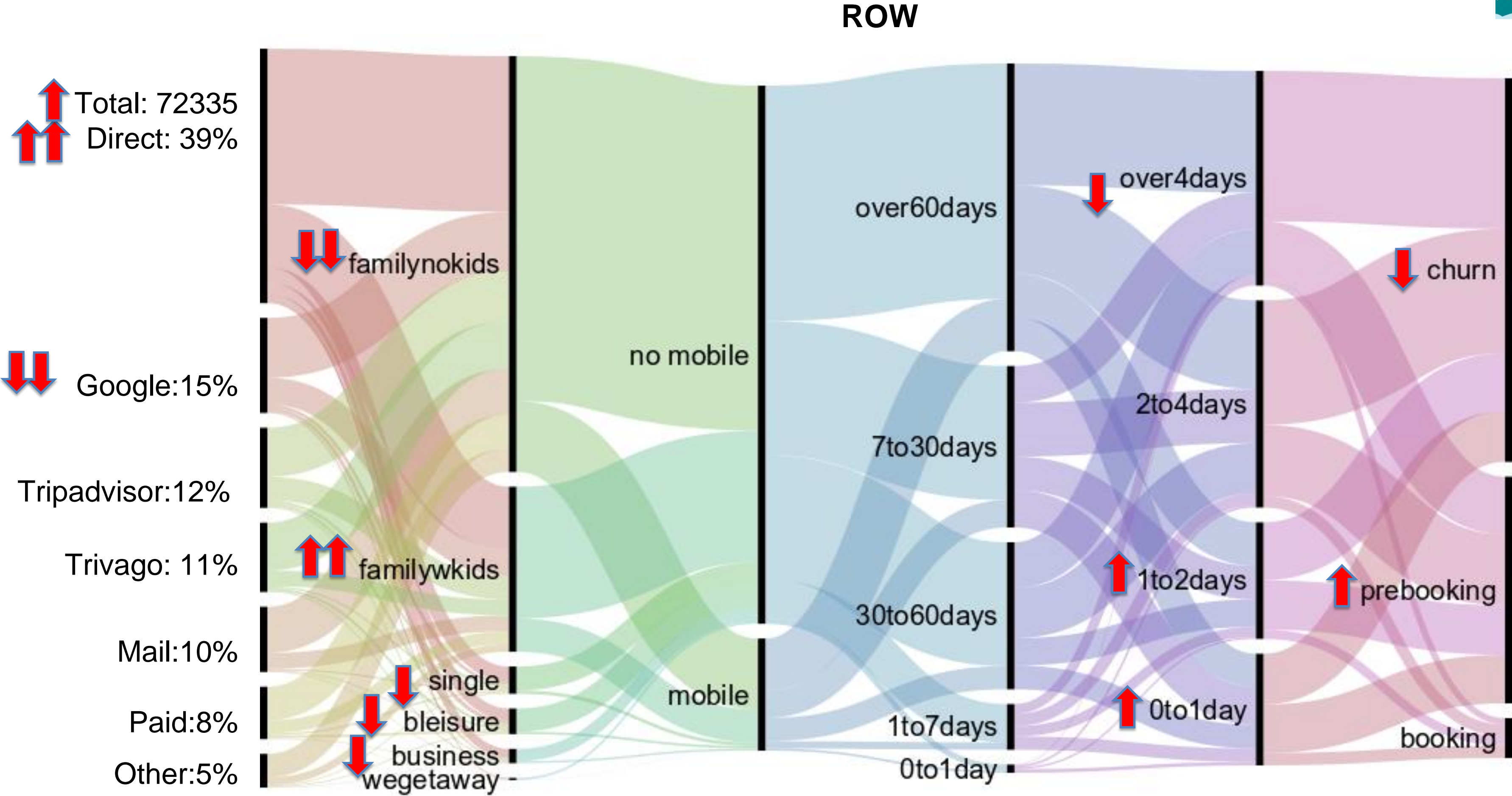
5. Seasonality effect



USA



5. Seasonality effect



5. Seasonality effect

