Do podcast and music compete with one another? Understanding users' audio streaming habits

Ang Li University of Pittsburgh

Alice Y Wang, Zahra Nazari, Praveen Chandar, Benjamin Carterette, **Spotify**

Interned 2019 summer with **Spotify Tech Research**

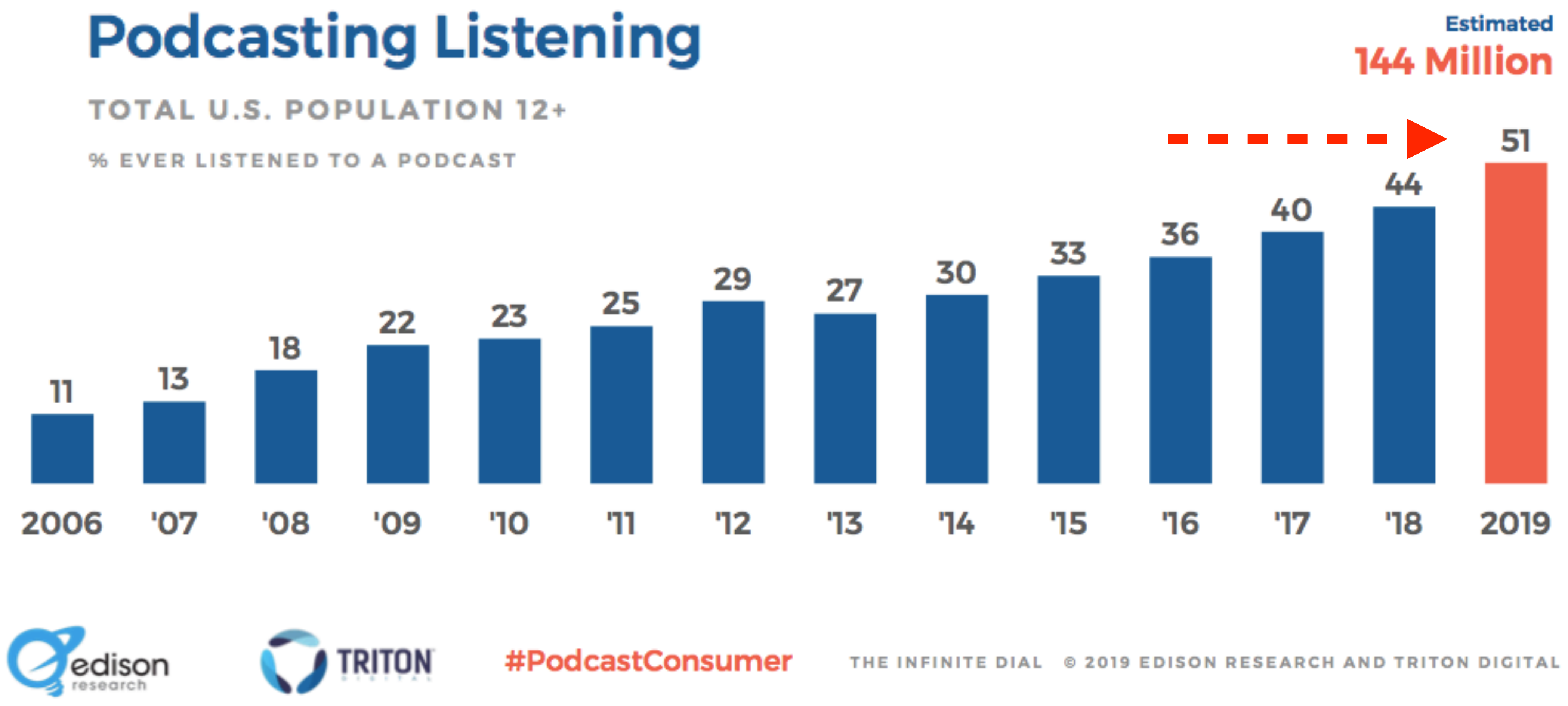
OUTLINE

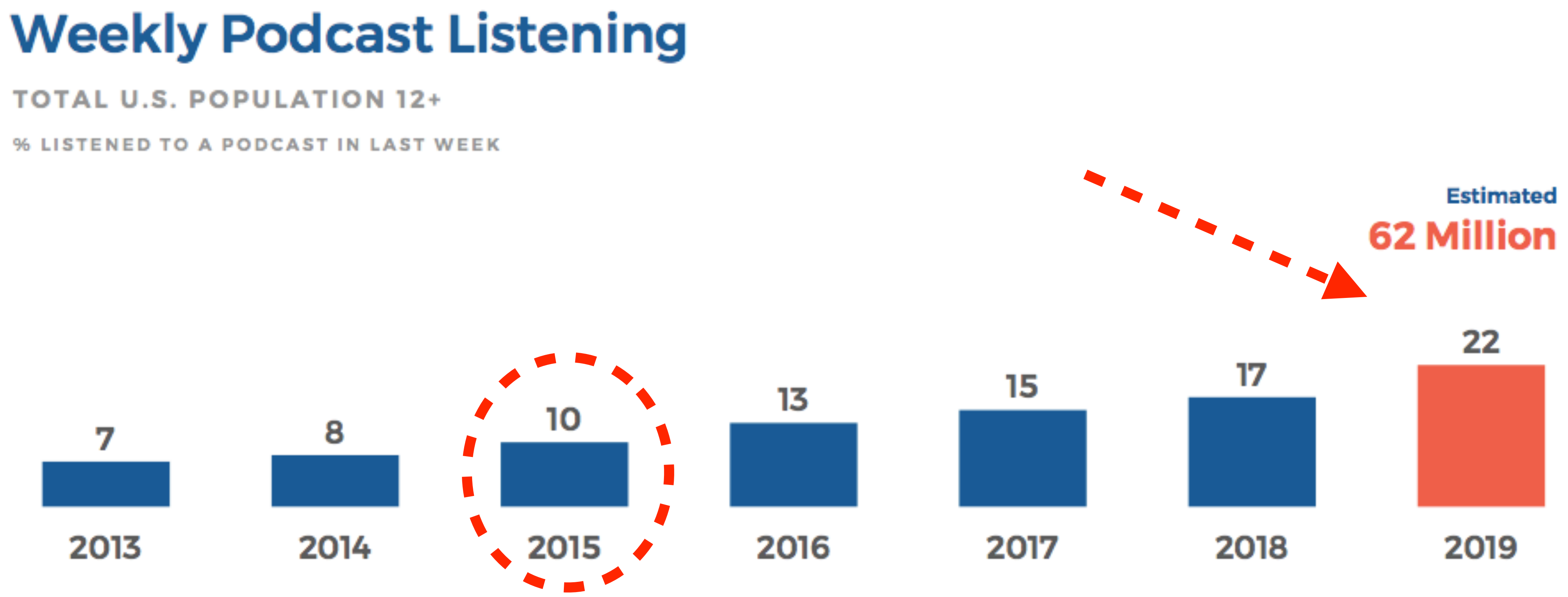
- Motivation and research questions
- Methodology
- Results
- Conclusion

Podcasts have become one of the fastest growing online audio streaming media

- **Podcast**: portable and on-demand forms of spoken-word audio content

- As of April 2019, there are over 29M episodes of Podcasts





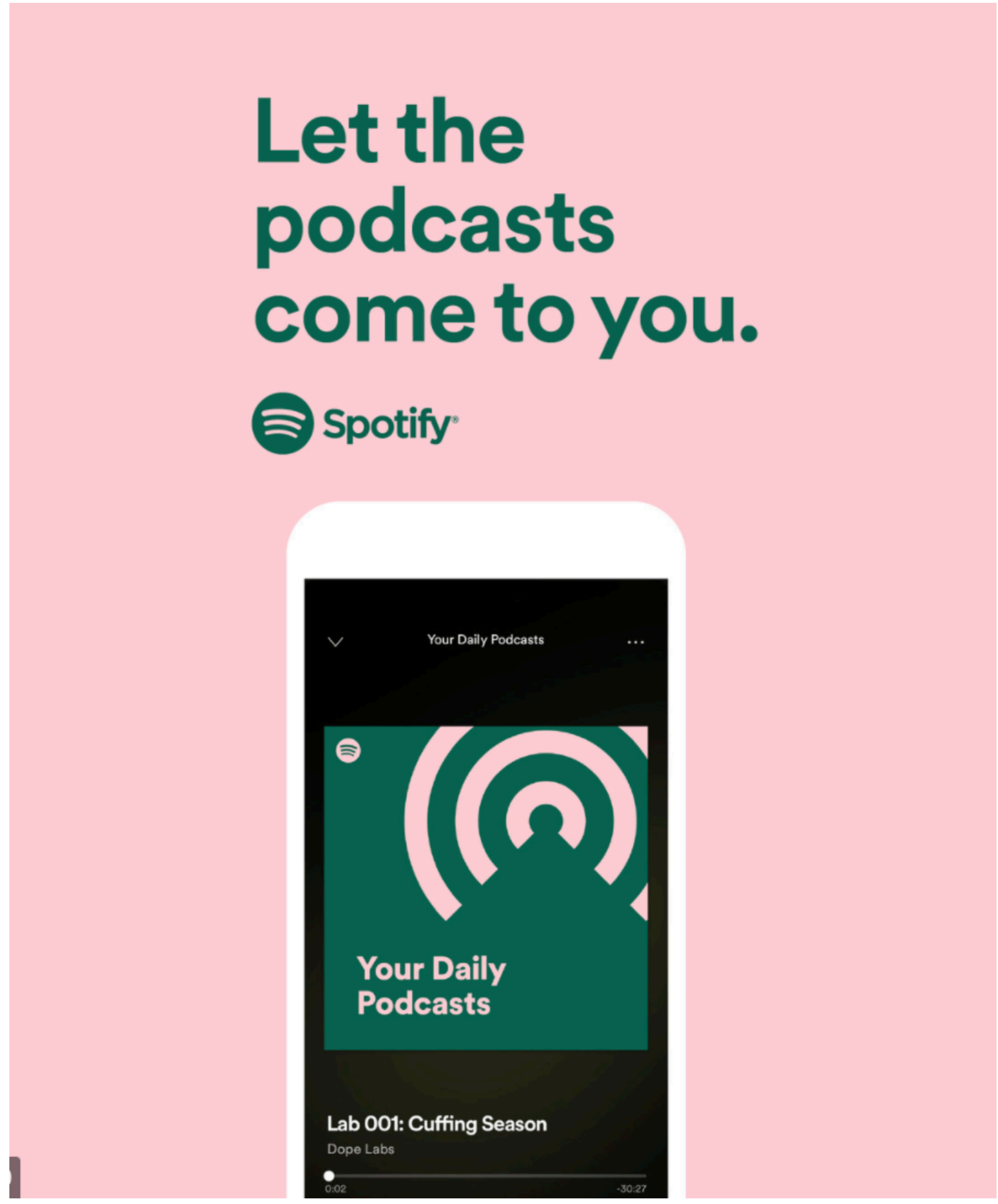
[1]. https://www.edisonresearch.com/the-podcast-consumer-2019/

The popularity of **podcasts** has created **great demand**

Market as of 2018:

- > 525K active podcast shows, > 18.5 billion downloads

- Several **music-focused** platforms started to expand services by adding **podcast content** alongside music
  - ☑ Satisfy user demand by providing diverse content



[1] https://www.fastcompany.com/40563318/apples-podcasts-just-topped-50-billion-all-time-downloads-and-streams

Incorporating podcasts introduces new challenges for these platforms

- The functional use for podcasts and music listening may largely **overlap**
- Two types of content may **compete** for the limited amount of time that users may allocate for daily audio streaming.

Music: entertainment, pleasure, passing time, education, facilitate social interactions

Podcasts: entertainment, relaxation, education.
One of the most popular topics among podcast consumers is **music**

☑ As a result, incorporating podcast listening to music streaming may influence and change the original way users consume music

Necessary for these platforms to understand the influence and users' listening habits

- It is necessary for platforms to understand:
  - ☆ What is the effects of injecting new type of audio content to users' listening habits
  - ☆ How users' listening habits change before vs. after the arrival of new content
- … provide best support for their users
  - ☑ Content consumers: support them with a better recommender system i.e. provide the right content at the right time
  - ☑ Content creators: make their shows more tailored based on users listening habits

IN THE CURRENT STUDY …

We focus on music listeners who incorporate podcasts to their listening activities for the **first time**, and ask:

RQ1. Would users' music listening habits change ***as a result of*** adopting podcast listening?

RQ2. a. For music users who have adopted podcast listening for the first time, what are their **listening habits** for podcasts vs. music? **b.** What are the **content** of podcasts they consume, especially *before the adoption*?

RQ3. When a user starts a new listening session, can we **predict** whether s/he will listen to a podcast based on their listening habits?

USING Spotify AS OUR RESEARCH PLATFORM

- One of the most popular (271M users) online audio streaming platforms
- Traditionally focus on music streaming
- Started to provide Podcast content in recent years

Define 3 **levels** of podcast engagement [based on internal user study]:

> 1 hour for a Podcast show
> 3 episodes of that show

Music Only

Only listened to music

Podcast Seekers

Just started listening to podcasts, still *explore*, *seek* Podcasts to listen to

Podcast Adopters

Have passed a threshold level of streaming podcast, **a proxy** for assuming that they **start to incorporate podcast into their listening habits**

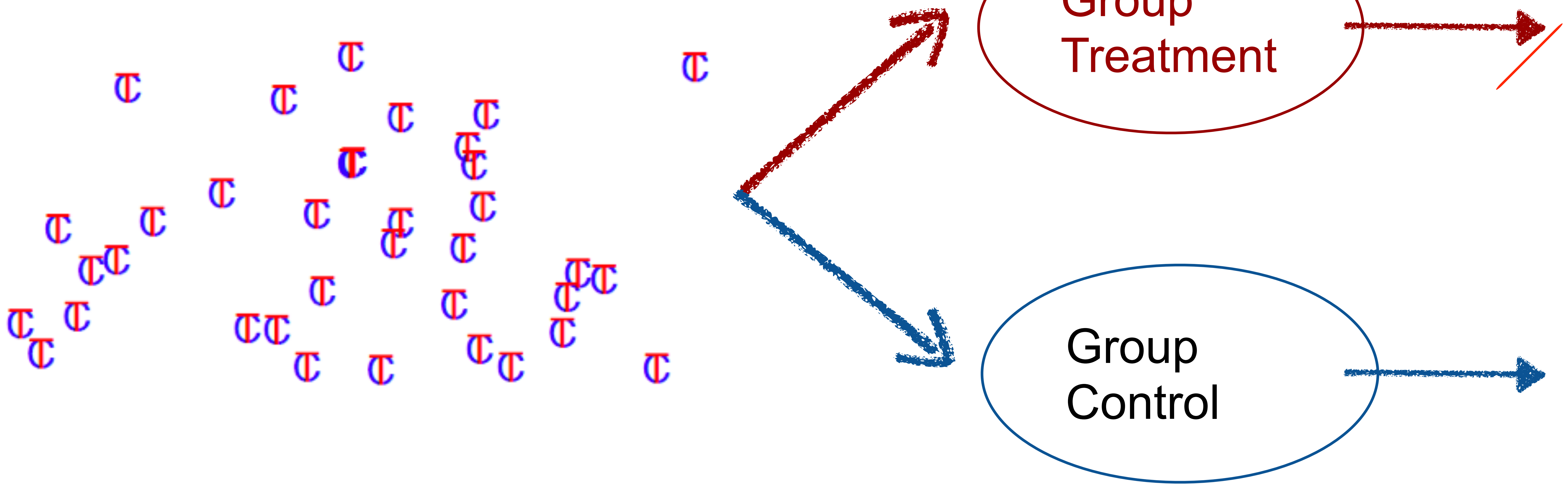
Adding podcasts may influence music listening?

Pairs of similar users

Controls & **Treatments**

- **similar** P(treated)

Received treatment



Outcome of Group T

= Causal effect

Outcome of Group C

[1] Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference.

[2] Hüseyin Oktay, Brian J. Taylor, and David D. Jensen. 2010. Causal discovery in social media using quasi-experimental designs. In Proceedings of the First Workshop on Social Media Analytics (SOMA '10). ACM, New York, NY, USA, 1-9. DOI=http://dx.doi.org/10.1145/1964858.1964859

PAIRS OF SIMILAR USERS: Music Only VS. Podcast Adopters

Pool of active users from May (US)

Treatment:
convert to Adopter

01/01 | 03/01 | 05/01 | 07/01

60 days | 60 days | 60 days

Adopters: Music | Music + **Pod Seeker** | Music + **Podcast Adopter**

Treatment users:
270K users

Music only: Music | Music | Music

Control users:
Never listen to Podcast

RQ1: Would users' music listening habits change ***as a result of*** adopting podcast listening?

Podcast Adopter:
RQ2a. What are their **listening habits** for podcasts vs. music?
– When do they listen to What type of content (i.e. podcasts vs. music)?

PAIRS OF SIMILAR USERS: Seekers VS. Adopters

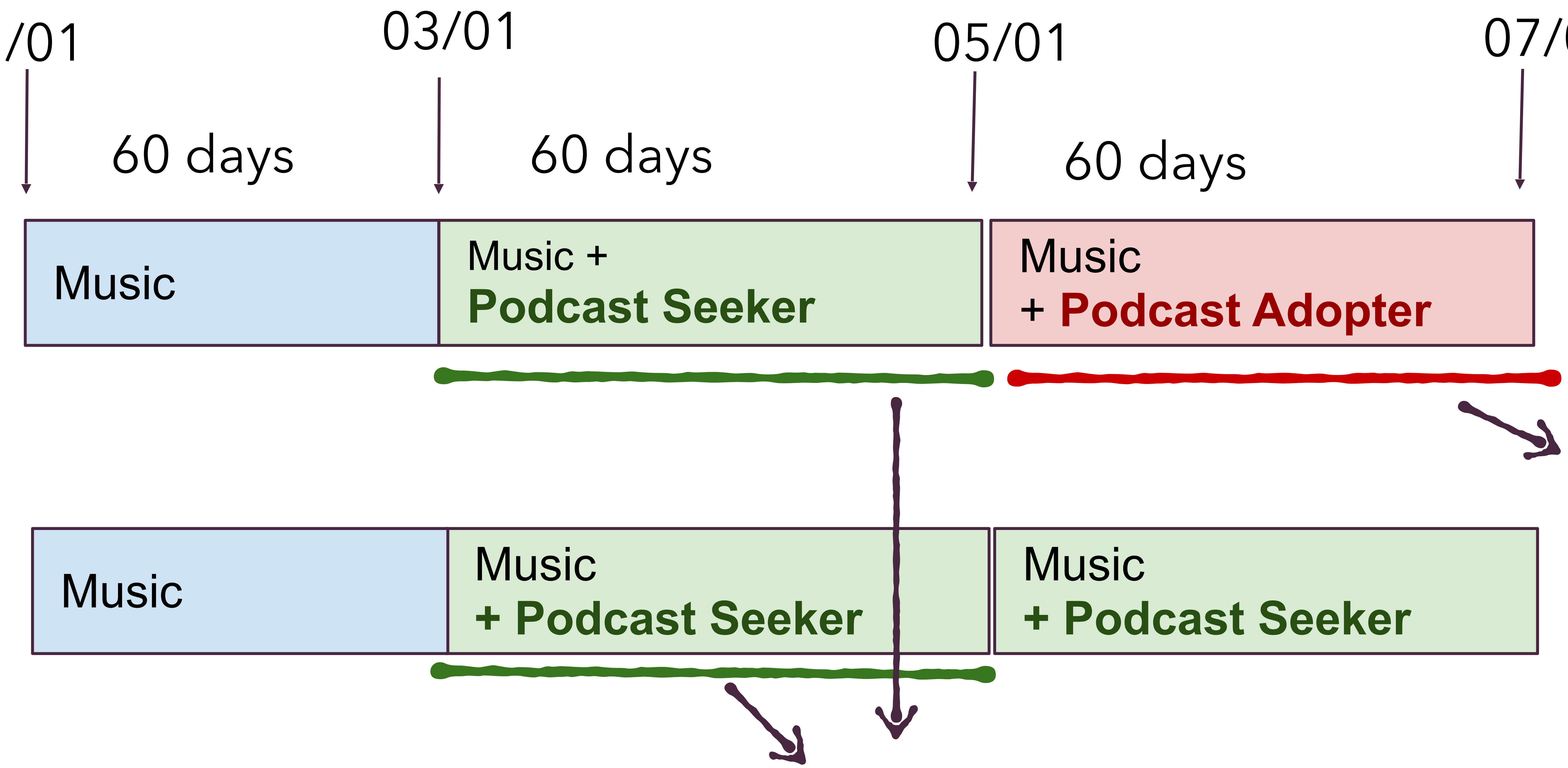
Pool of active users from May (US)

Adopters

Treatment users:
270K users

Seekers

Control users:
Never converted to adopters

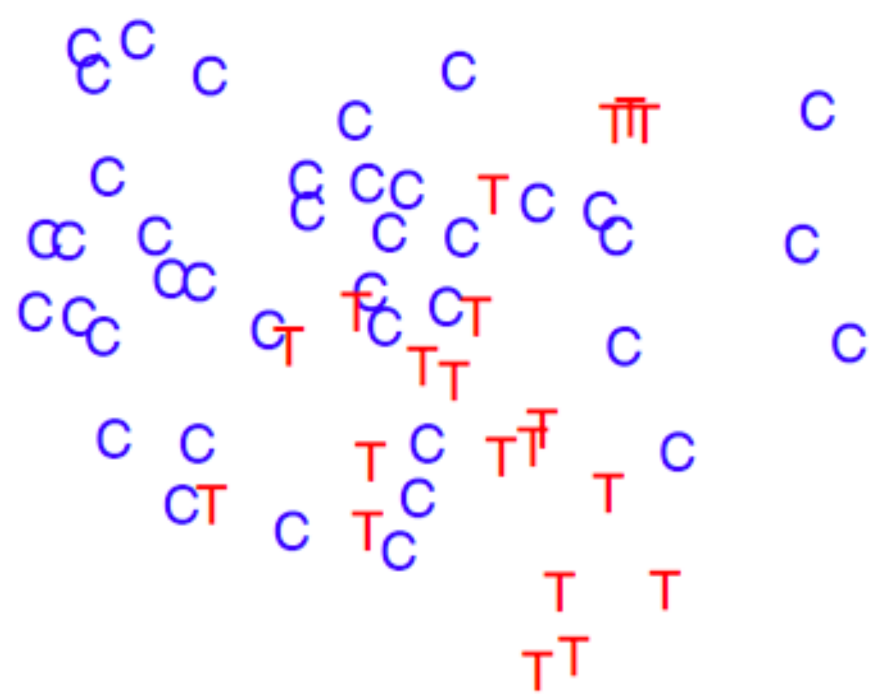


Podcast Adopter:
RQ3. Predict: when a user launch a new session, whether he/she will listen to Podcast?

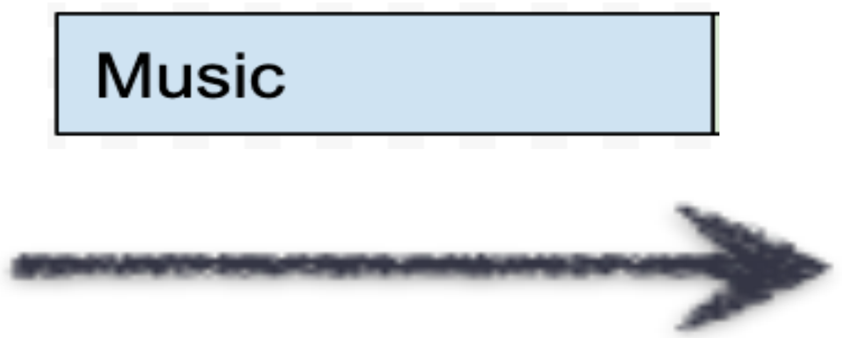
Examining the **seeking period** for both groups:
RQ2b. What are the **content** of podcasts they consume, especially **before the adoption**? – differences on listened show types for adopters vs. seekers?

PROPENSITY SCORE MATCHING TO FIND SIMILAR PAIRS

Treatment
Control candidates



Music listening behavior before listen to Podcast

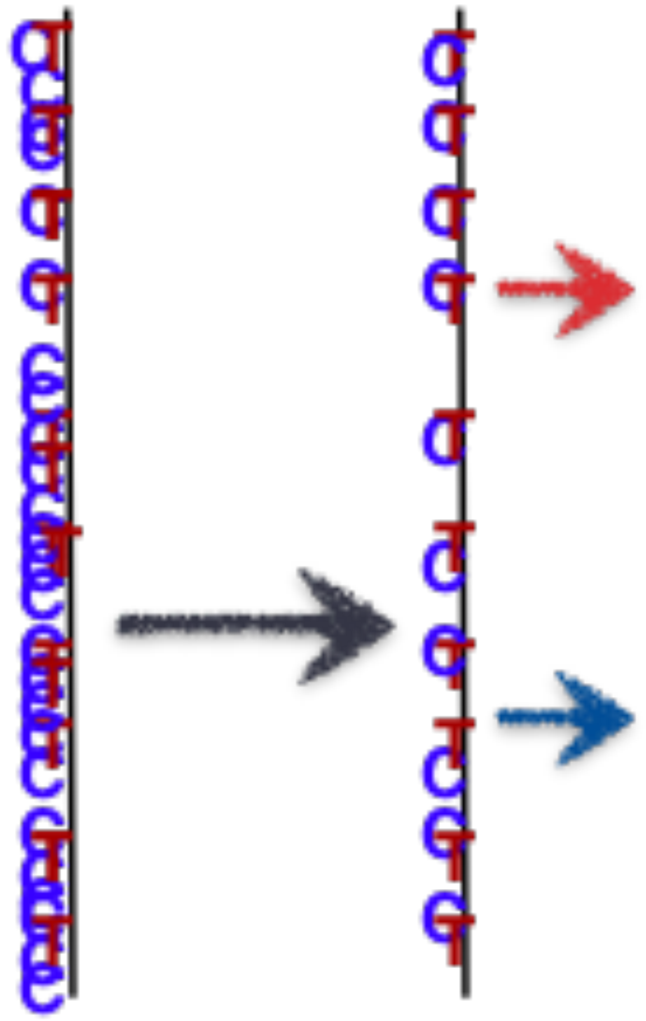


16 Confounds Xi:

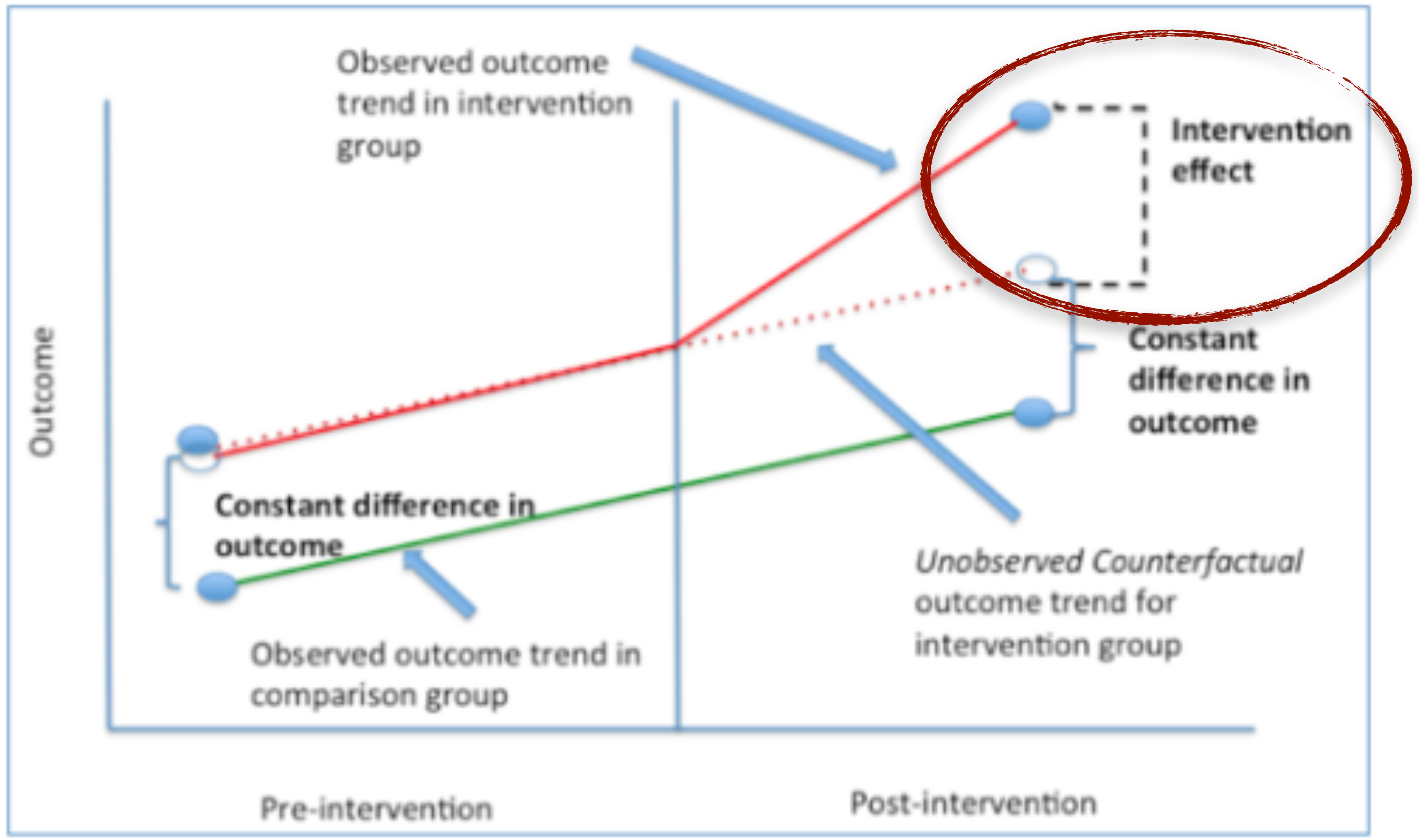
Gender, age
Registration
Music streaming behavior:
- Overall - streams/time
- Dayparts
- Weekdays/Weekend
- # of artists
- # of tracks

Estimate Propensity score:

$$\pi_i := \pi(X_i) = Pr(T_i = 1 \mid X_i).$$



Match



Difference-in-difference

= ***changes*** in <u>outcomes</u> over time between the intervention group and the control group

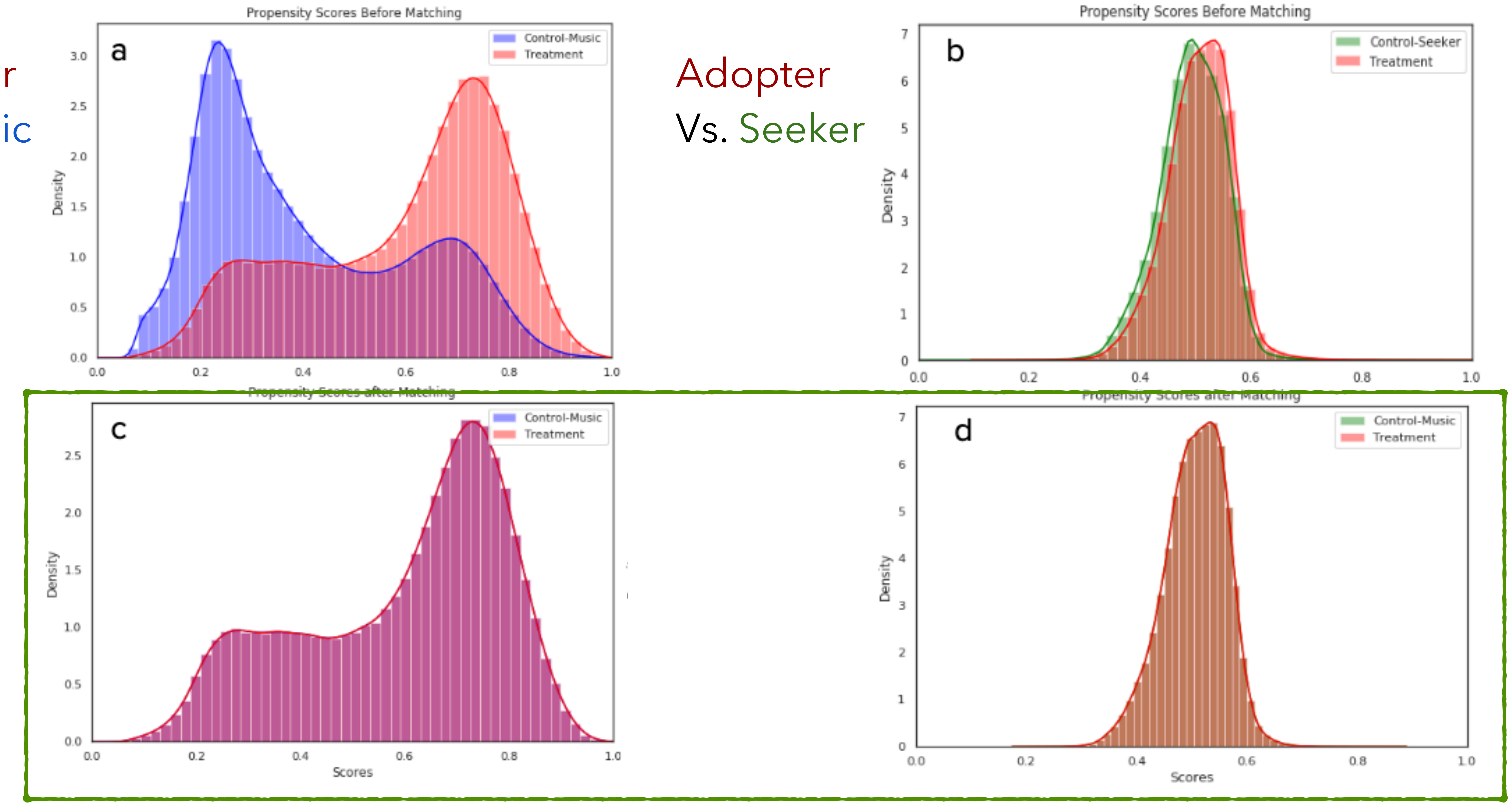
[1] Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, *84*(1), 151-161.
[2] Donald, S. G., & Lang, K. (2007). Inference with difference-in-differences and other panel data. The review of Economics and Statistics, 89(2), 221-233
[3] Difference in difference estimation: https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation

PROPENSITY SCORE: Before VS. After MATCHING

- We are able to match each treatment user with a control user based on their *propensity scores*
- After matching, the distribution of *propensity scores* for treatment and control groups overlapped

Adopter
Vs. Music

Adopter
Vs. Seeker



Matching:
270K treatment
270K control

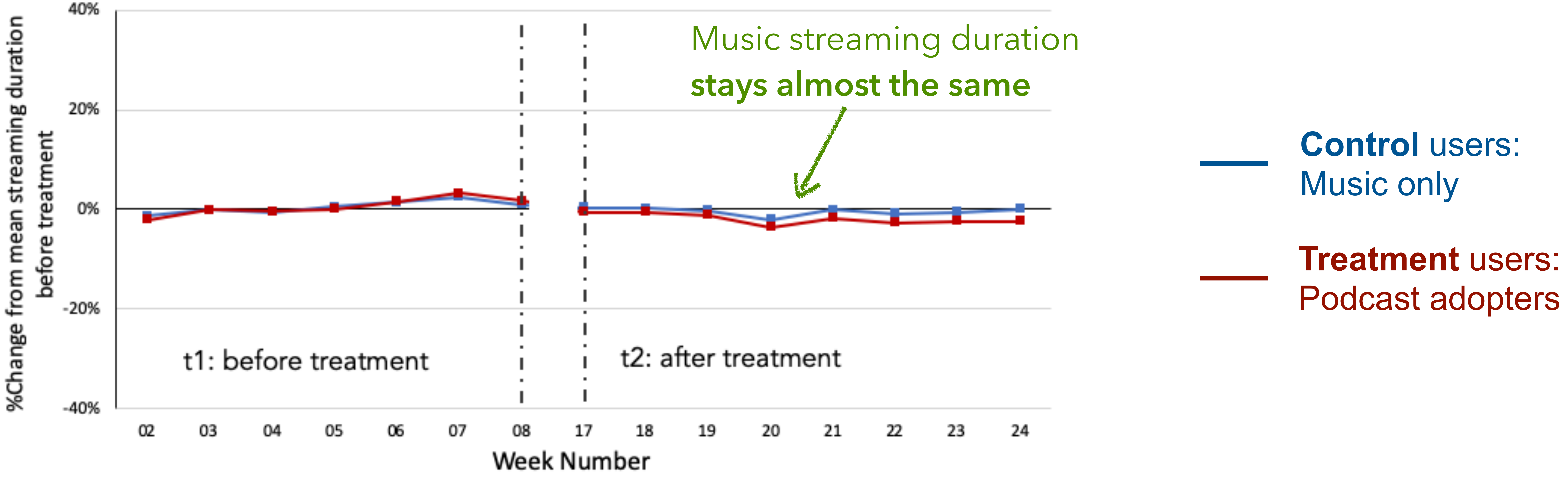
Music Only Vs. Podcast Adopters

RQ1: Will users' **Music** *listening habits* change as a result of adding **Podcast** listening?

Users music listening habits stay almost the same

Music streaming duration:

- After podcasts were adopted, users spent just 1% shorter time (than before) listening to music

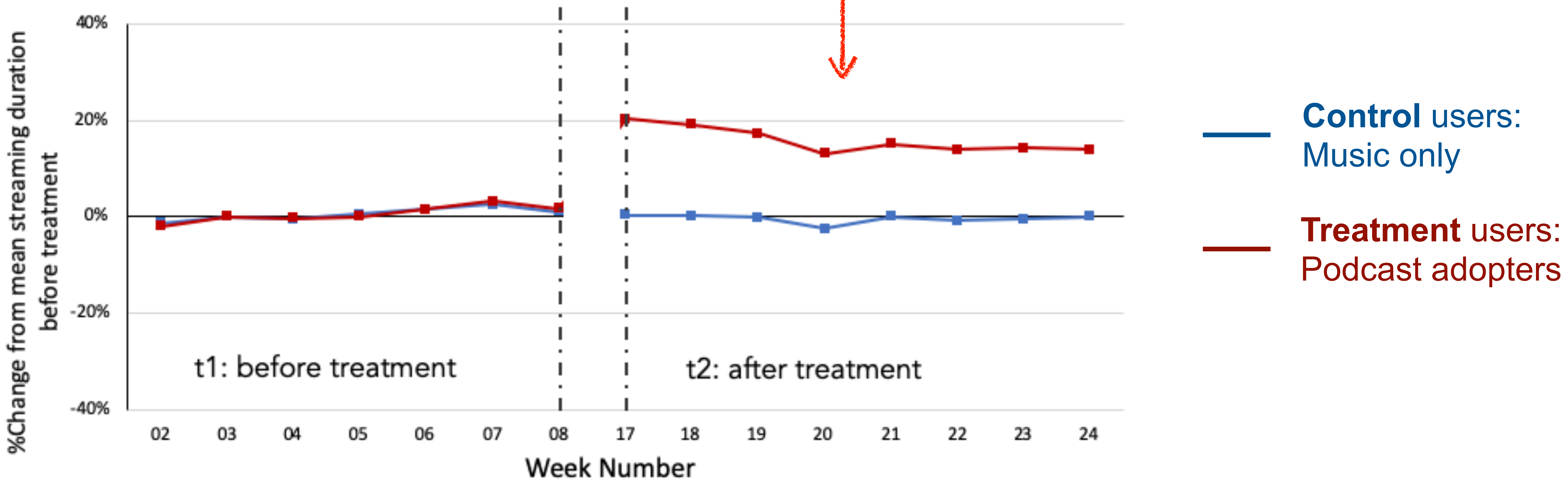


- Had similar findings for **music listening frequency**, please refer to paper for details

Users add additional time in listening podcasts

Total streaming duration:

- After podcasts were adopted, users spent **20% (at most) longer time** streaming in platform than before

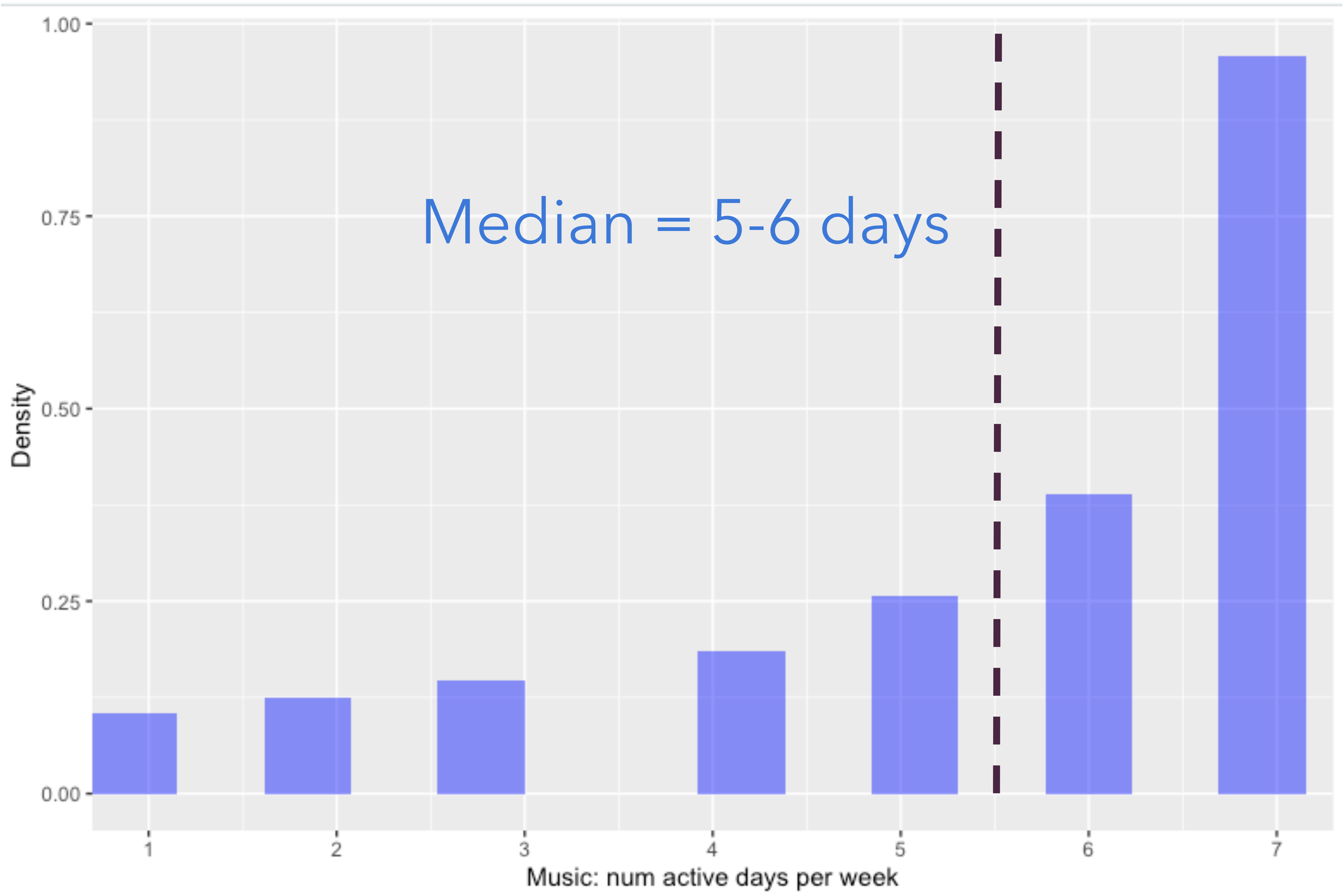


RQ2a:

Podcast🎙 Vs. Music🎧 both play unique role – very different *listening habits*

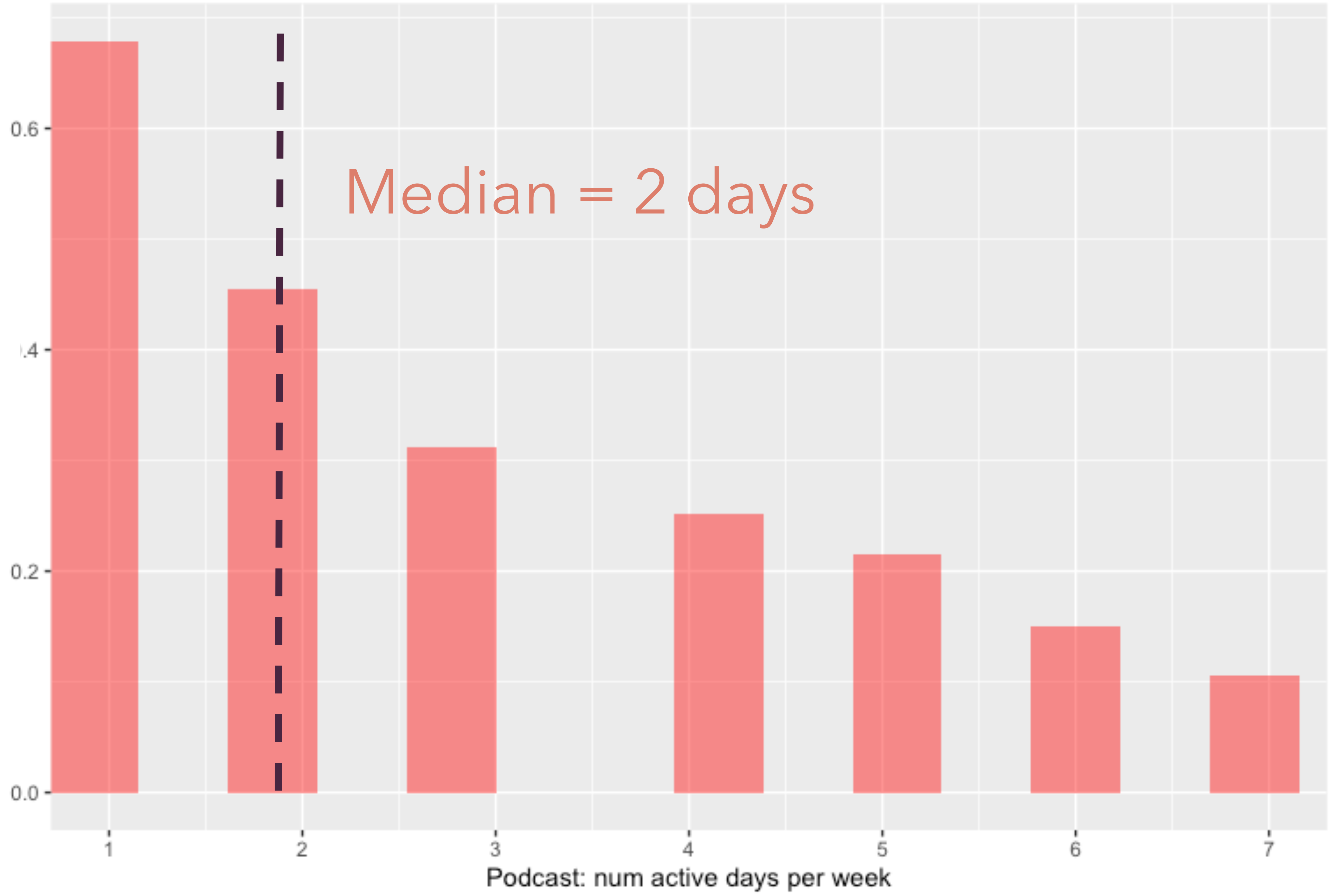
Music listening is a daily activity; Vs. Podcast listening is a weekly activity

Weekly: # of active days listened to **music**



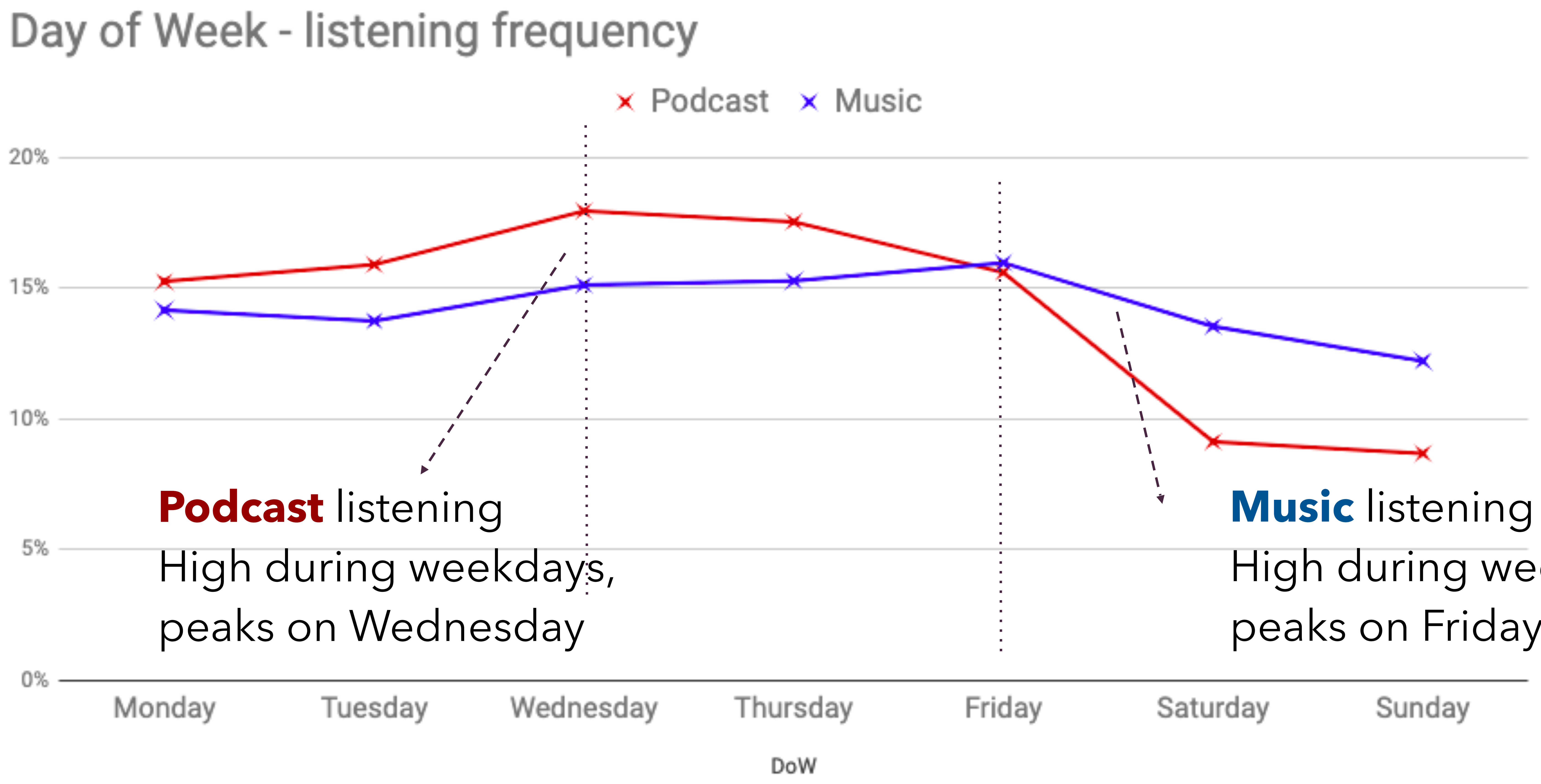
Music listening is a daily activity
In average, **5 -6 Music active days** per week

Weekly: # of days listened to **Podcast**



Podcast listening is a weekly activity
In average, **2 Podcast active days** per week

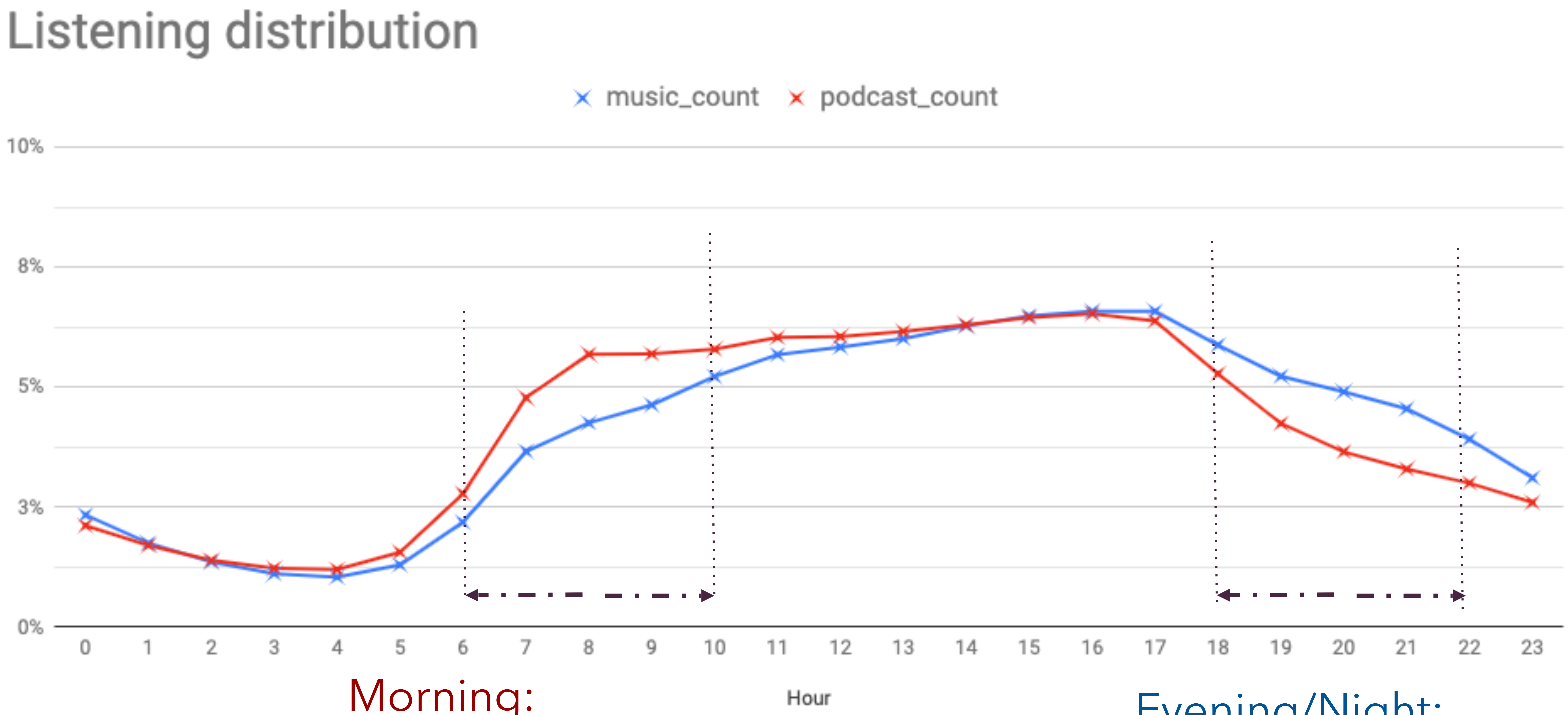
Podcast Adopters: Listening frequency across a Week



Podcast listening
High during weekdays,
peaks on Wednesday

Music listening
High during weekends,
peaks on Friday

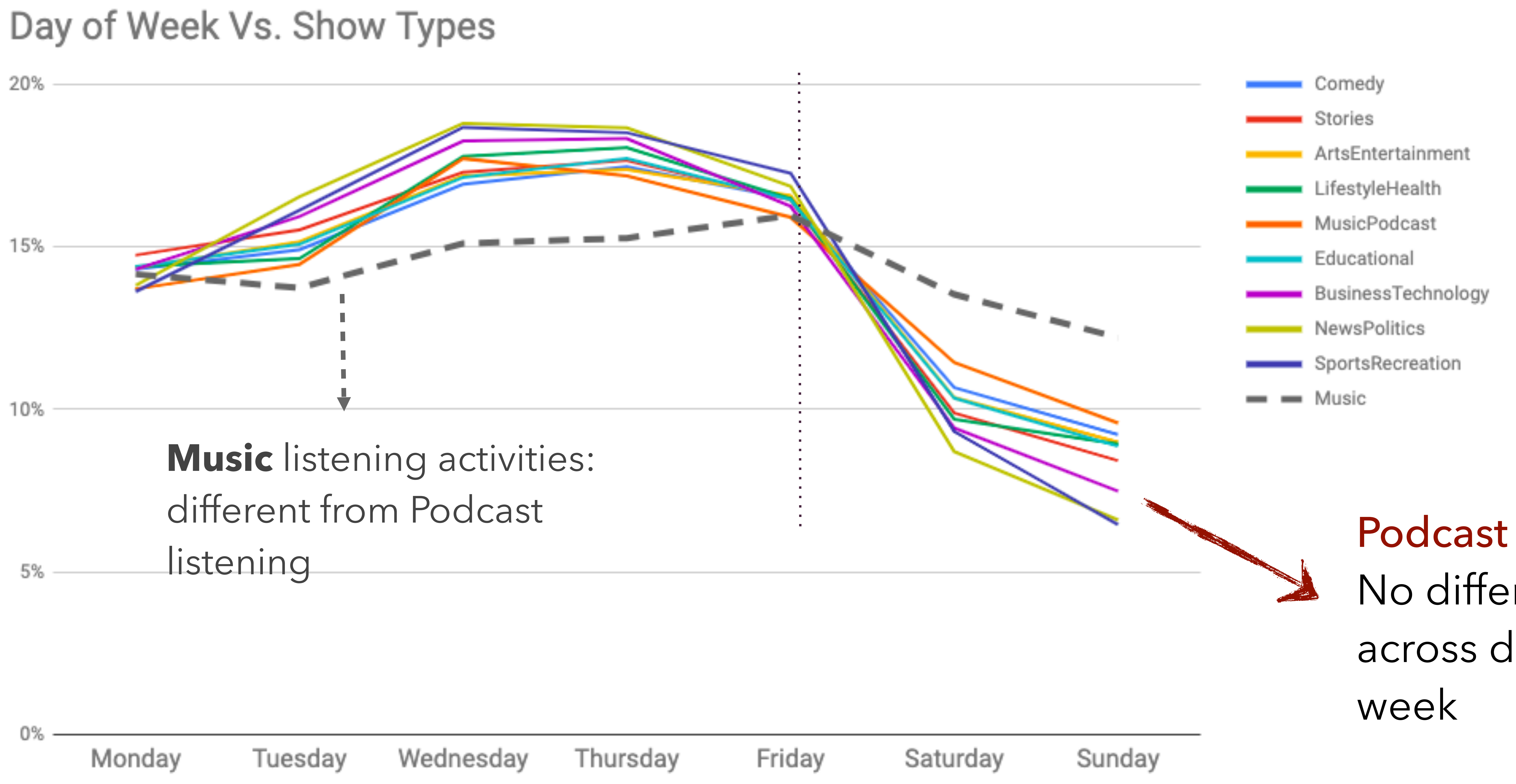
Podcast Adopters: Listening frequency across a Day



Morning:
Podcast listening time

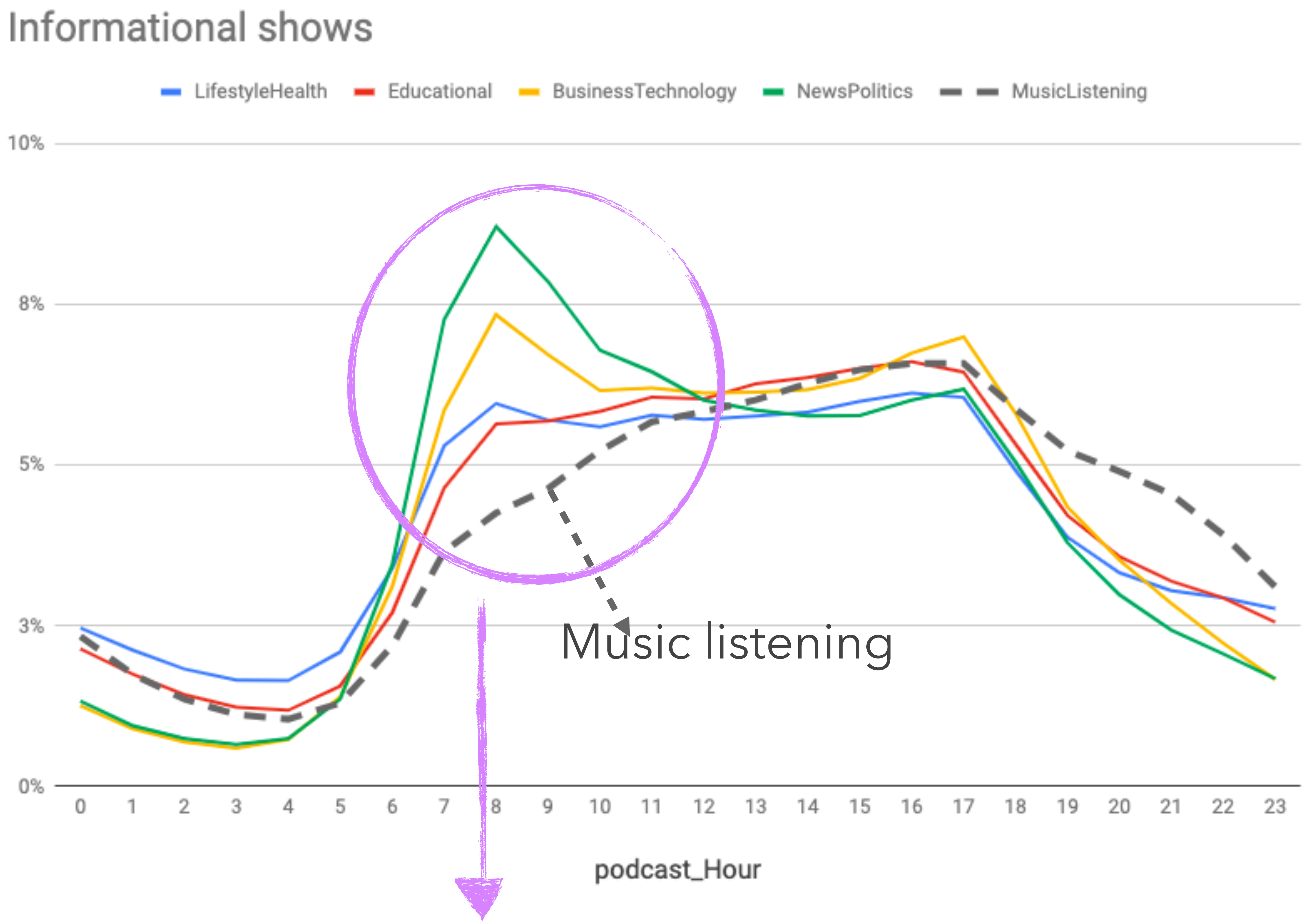
Evening/Night:
Music listening time

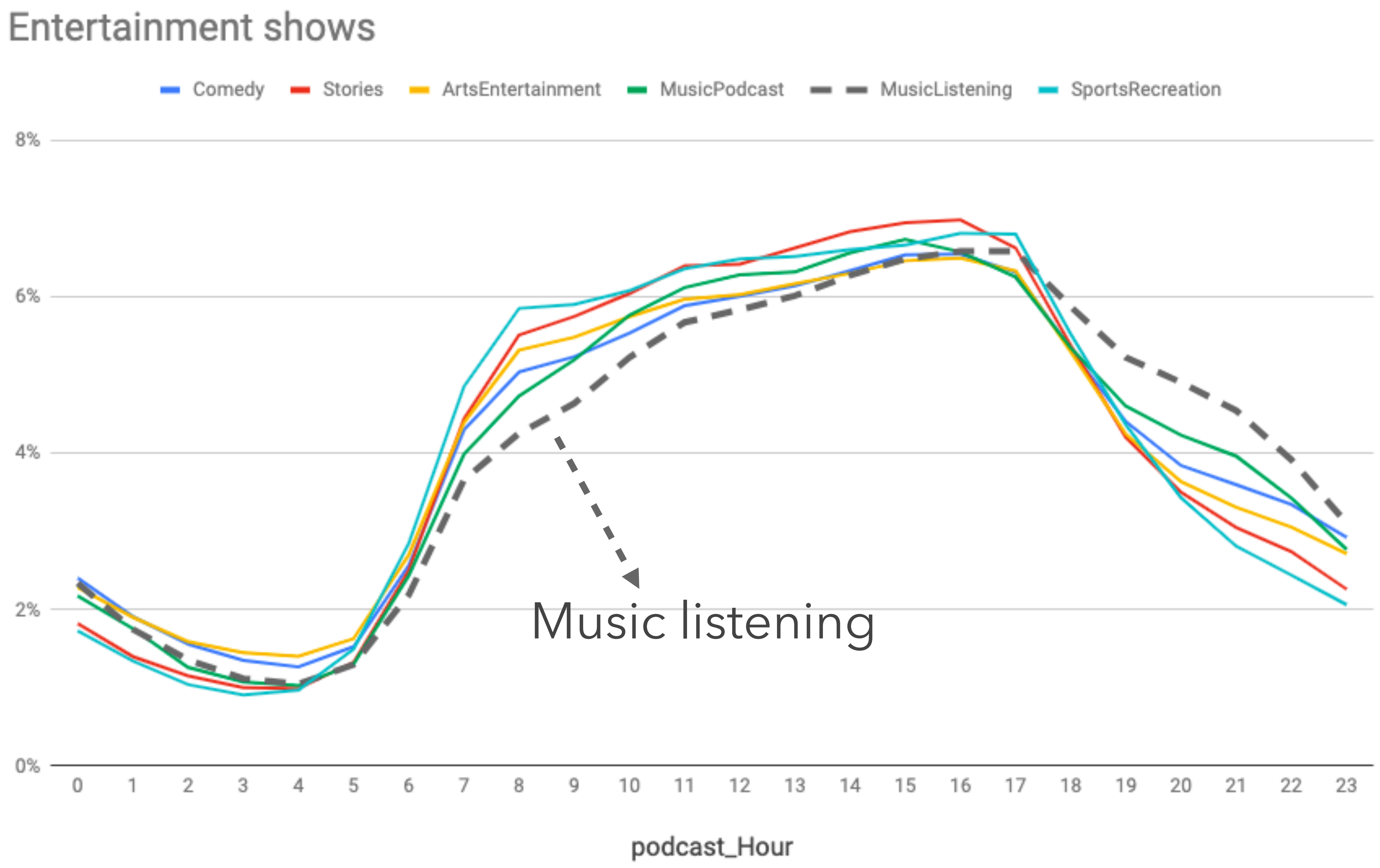
Podcast Adopters: Different shows Vs. Day of Week



Podcast listening:
No difference for topics across different day of a week

Podcast Adopters: When do they listen to different shows?





Compare to **Music listening**
→ **Informational** show streaming peaks on early Morning

Entertainment show streamings' trend is more similar to **Music listening**

Podcast Seekers Vs. Podcast Adopters

RQ2b: Consumed Podcast Content:

Compared to seekers, what are the podcast show types that adopters consumed, especially before the conversion?

Shows watched during seeking period for both groups

Compared to **seekers,**

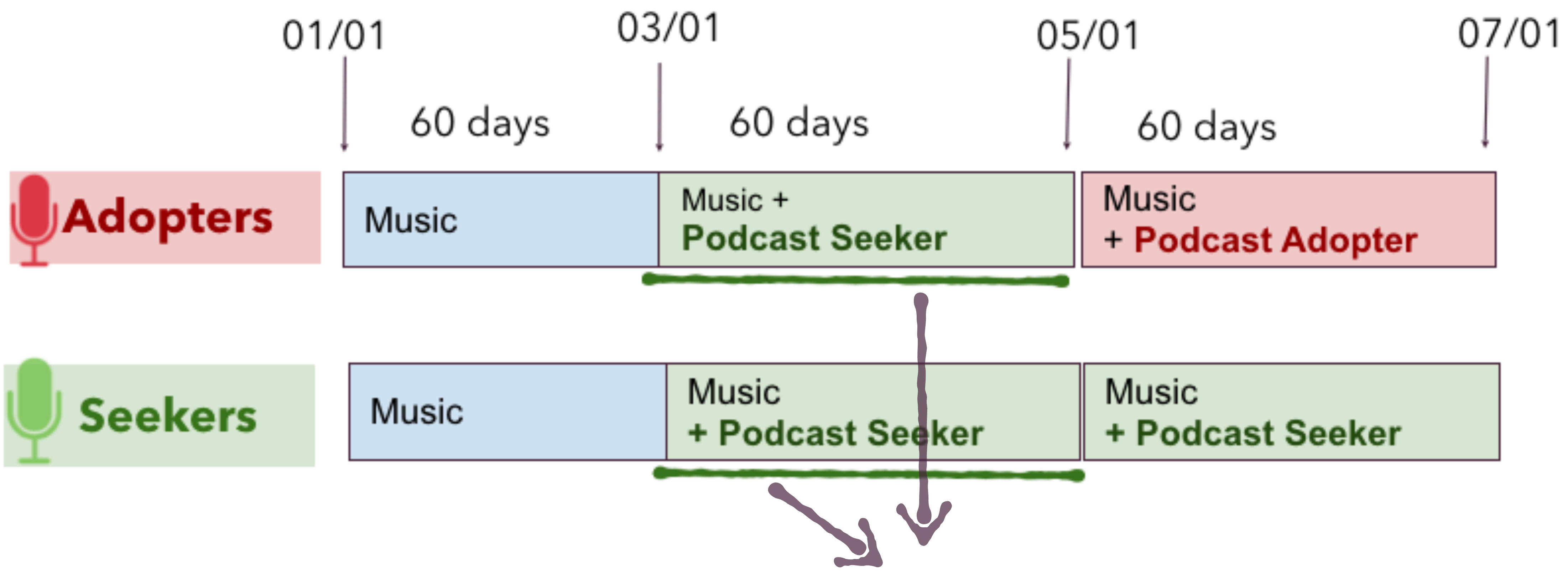
Adopters tend to

Listen significantly **more**

- Sports, Stories, Comedy, True Crime

Listen **less**

- Music, Art, Health
- Educational, News, Business



Seeking period:
Differences in the watched show types?

The streamed show types in users' podcasts seeking period can be very important – there are certain show types are more attractive to users who have adopted podcast listening

The streamed show types during seeking period are strong predictors for conversion of Podcast adopters

- Retrieve only the **first** day activities of podcast listening for users from two groups (seekers, adopters)
- Predict a user's membership – who will become **podcast adopters** eventually?

Features: extracted based on users' first day activities of Podcast listening

- **Show types**: # of Stories, Comedy, True Crime, Music, Education, etc
- → *Referral* types: (the stream is referred from) a browse, search, home, library
- → *Activity*: # streams, # shows, # episodes

Data: ~500K users from two groups (seekers, adopters)

- Excluded users who completed conversion within their first day
- Training data: 70%; Test data: 30%

Show types are important in all predictions

- Logistic regression in predicting user's membership
- Iteratively add features and check the feature prediction power

Features	Training model accuracy	Testing model accuracy	**Top 3 Predictive Features**
Show types	**69.50%**	**68.80%**	1. # of listened Sports/Recreation shows 2. # of listened True Crime shows; 3. # of listened Comedy shows
+ Referral types	70.40%	70.30%	1.# of listened Sports Recreation shows; 2.# of listened True Crime shows; 3.# of referral from library
+ Activity	71.00%	70.90%	1.# of listened Sports Recreation shows; 2.# of listened True Crime shows; 3.# of referral from library

- Including ONLY Show types features, the model can already achieve the accuracy rate as 68.8%.
- Adding referral types features and users' activities features, the model improves only 2%
- Across all three models, the show types features have consistently been identified as top predictors

SO FAR WE UNDERSTAND …

Although there is a **mild competition**, Podcast and Music are both important and do not substitute for one another: <u>users open a new time window listen to Podcasts</u>

Daily activity: 5-6 active days per week

Entertainment:

- More likely to happen during Evening
- On weekends, Peak on Friday

Weekly activity: 1-2 active days per week

Information/Education:

- More likely to happen during Early morning, for informational types of shows
- On weekdays, Peak on Wednesday

Consumed Podcast Content during Seeking period is important:

There are certain **show types** that have lower entry barriers for Podcast listening – **show types** are strong predictors of the user conversion

RQ3: Prediction:
When a user starts a new listening session, will she/he listen to a podcast?

SESSION LISTENING PREDICTION

Session Data:

- A session: idle time > 10 mins
- 10% positive sessions (have Podcast listening)

Features (X):

User features:
Gender, Age,
Registration

Music and Podcast
listening activities:

The **last 1 session**,
So far for **today**
So far for **this week**

The **current Session :**
Time, day of week

Target (Y):

Current session
- Podcast listening?

Unbalanced:
10% Podcast
listening session

WILL USER LISTEN TO PODCAST IN THIS SESSION?

5 folds cross validation **at the user level:**

- Make sure that sessions generated from one user will not appear in different folders
- Avoid potential issues such as predicting past events based on the future

	Logistic regression [mean, std]	**Random Forest** [mean, std]
User features + the last 1 session	**Train_f1: [0.363, 0.0135]** **Test_f1: [0.361, 0.0191]**	Train_f1: [0.536, 0.0141] Test_f1: [0.534, 0.0323]
+ current time	Train_f1: [0.404, 0.0134] Test_f1: [0.402, 0.0206]	Train_f1: [0.563, 0.0143] Test_f1: [0.561, 0.0343]
+ so far for today	Train_f1: [0.413, 0.0123] Test_f1: [0.411, 0.0258]	Train_f1: [0.590, 0.0195] Test_f1: [0.589, 0.0370]
+ so far for this week	Train_f1: [0.439, 0.0130] Test_f1: [0.437, 0.0301]	**Train_f1: [0.601, 0.012]** **Test_f1: [0.599, 0.026]**

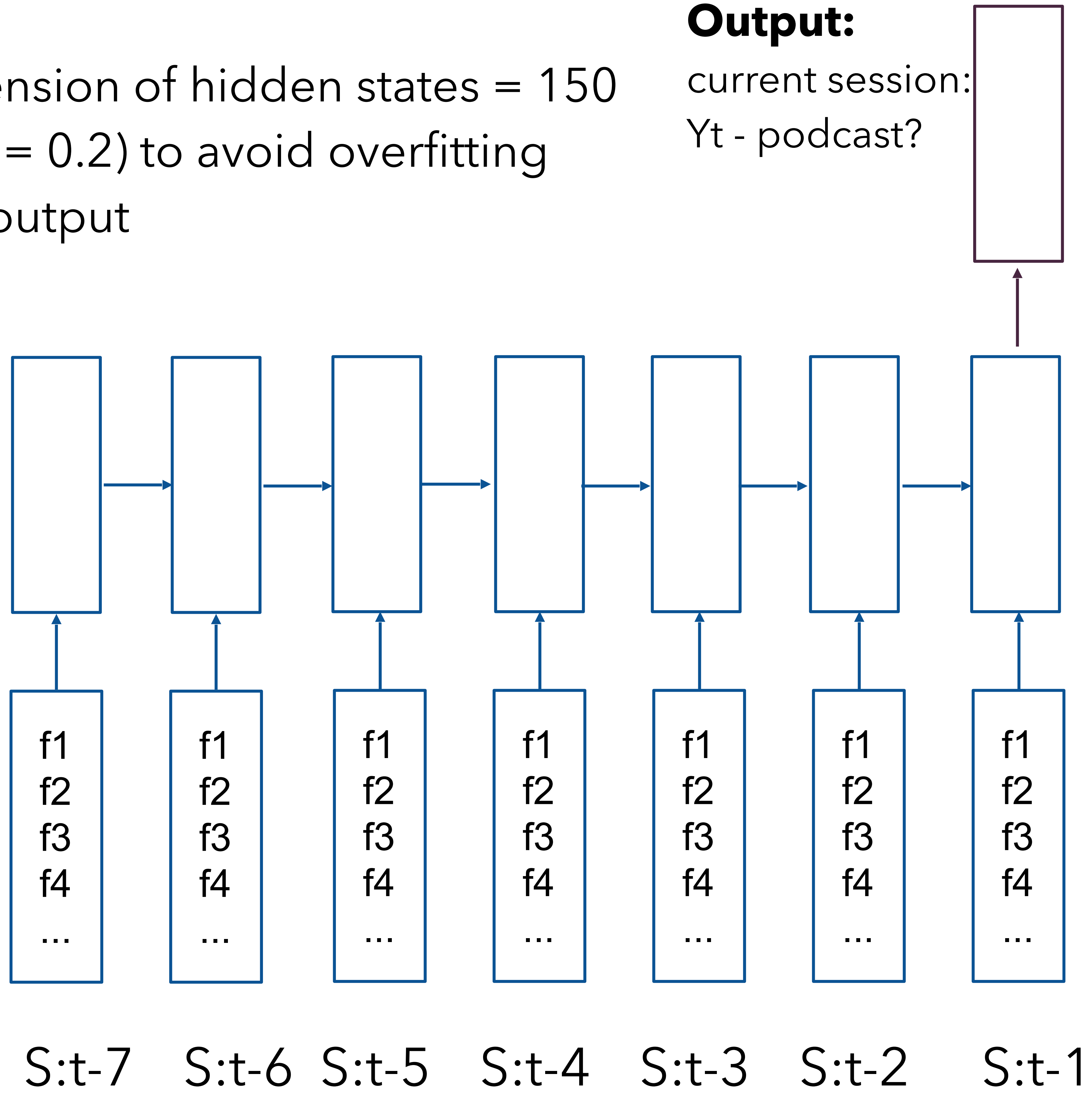
- Logistic regression with [User features + the last 1 session] as a baseline model
- Iteratively add features and check the model performance, best model has mean F1= 0.6

LSTM using past 7 sessions demonstrates better performance

Model:

1. LSTM layer with dimension of hidden states = 150
2. A dropout layer (rate = 0.2) to avoid overfitting
3. Dense layer for final output

Output:
current session:
Yt - podcast?



Input of S: t-n

User

Current session

- time
- listening

	Train [mean, std]	Test [mean, std]
Recall	[0.671, 0.021]	[0.666, 0.032]
Precision	[0.828, 0.013]	[0.825, 0.027]
F1	**[0.741, 0.011]**	**[0.737, 0.027]**
Accuracy	[0.963, 0.001]	[0.962, 0.002]

* Average time for the past 7 sessions is about 24 hours

Conclusion

WHAT WE STILL DON'T KNOW …

1. **Causal inference:** only accounted for **observable** confounders; **Things may happen outside Spotify**
2. Our observation window is **only 2 months** after conversion – the observed behavior maybe related to the **novelty of the experiences**
3. Our **definition** for "Podcast Adopters" is based on a threshold level of user engagement with Podcast – there can be potential selection bias
4. The listening habits discovered in the current study may be **caused by inherent differences in the two media forms** i.e. users listen to musics repeatedly but listen to podcasts only when new episodes are released.

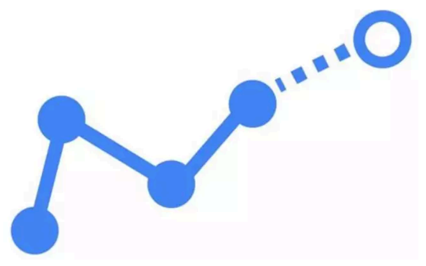
MAIN TAKEAWAYS:

Podcast and Music are NOT substitutable with one another:

- Users open another time window (20% longer streaming time per week) to listen to podcasts.

Podcast and Music both play important and unique roles:

- Users demonstrate different listening *habits*
  - Music: daily activities, during evening/night, weekends
  - Podcast: weekly activities, during morning, weekdays, and for information
  - The Podcast Content consumed during **Seeking** period is important: There are certain show types that have lower entry barriers for Podcast listening



Finally, using the above results to create input features to a machine learning model, a podcast listening session is predictable with high accuracy rate

Spotify

Thanks!
Questions?

Spotify®