Unveiling social desirability scales by comparing individuals' responses to an online survey with their streaming history data

Thomas LOUAIL
CNRS researcher at Géographie-cités & PACTE

in collaboration with Yann RENISIO
CNRS researcher at CRIS (Sciences Po, Paris)

FORS Research and methods seminar
28/05/2024

Acknowledgements

The work presented here has been partly funded by the French national research agency (ANR) under the grant ANR-19-CE38-0013 ‘RECORDS’. It is a collaborative project involving three research units of CNRS Humanities and Social Sciences — Géographie-cités, Centre Marc Bloch and the Center for Research on Social Inequalities — Orange/SENSE and the streaming platform Deezer.

The work presented in following slides is the result of a pluridisciplinary collaboration with (by alphabetical order) Amélie Beaumont, Jean-Samuel Beuscart, Samuel Coavoux, Philippe Coulangeon, Robin Cura, Brenda Le Bigot, Manuel Moussallam, Yann Renisio and Camille Roth.

Other active participants to the RECORDS project include Myriam Boualami, Pierre Gallinari-Safar, Noé Latreille, Darick Lean, Marion Maisonobe, Kristina Matrosova, Alvin Opler, Anne-Cécile Ott, Anne-Sylvie Pharabod, Jérémie Poiroux, Bruno Massoni Sguerra and Dougal Shakespeare.

Context
Objectives of the RECORDS project

1. Re-thinking categories in sociology of cultural consumption
Switching from declarative data collected at the level of music genres ("I like rock and jazz")
to mixed, declarative+observational data (digital traces) of effective consumption,
collected at the level of tracks/albums/artists ("which rock and which jazz? how much? how long?")

2. Measuring the diversity consumed on streaming platforms
Evaluate the relevance of statistical models linking social properties and music preferences
that have been built on declarative data only

3. Measuring the effects of algorithmic recommendation
What are the measurable effects of a recommendation system among its users,
after several years of use/exposure?

4. Investigate the relationship between declared music preferences and music streaming practices
Measure over- and under-representations and potential influence of legitimacy in displaying tastes

Context
Objectives of the RECORDS project

1. Re-thinking categories in sociology of cultural consumption
Switching from declarative data collected at the level of music genres (“I like rock and jazz”)
to mixed, declarative+observational data (digital traces) of effective consumption,
collected at the level of tracks/albums/artists (“which rock and which jazz? how much? how long?”)

2. Measuring the diversity consumed on streaming platforms (S. Coavoux & P. Coulangeon)
Evaluate the relevance of statistical models linking social properties and music preferences
that have been built on declarative data only

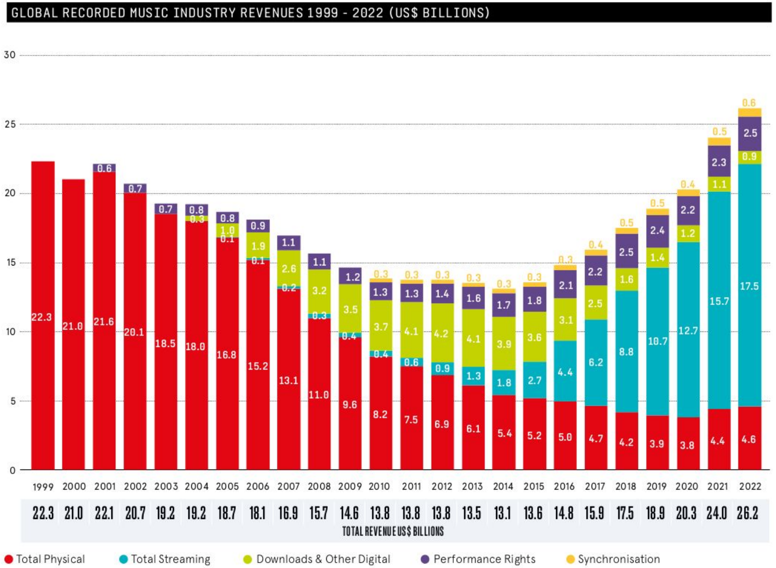
in this talk

3. Measuring the effects of algorithmic recommendation (C. Roth)
What are the measurable effects of a recommendation system among its users,
after several years of use/exposure?

4. Investigate the relationship between declared music preferences and music streaming practices
Measure over- and under-representations and potential influence of legitimacy in displaying tastes

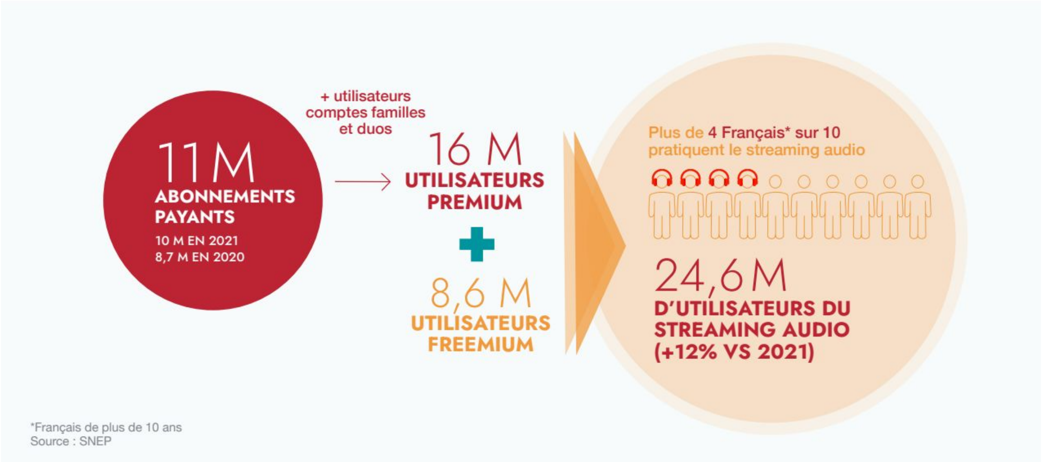
Context
The increasing importance of streaming in the global recorded music market

World



About 600M premium accounts users

France



deezer

16M active users worldwide
about half in France

3 to 5 million premium account users in France

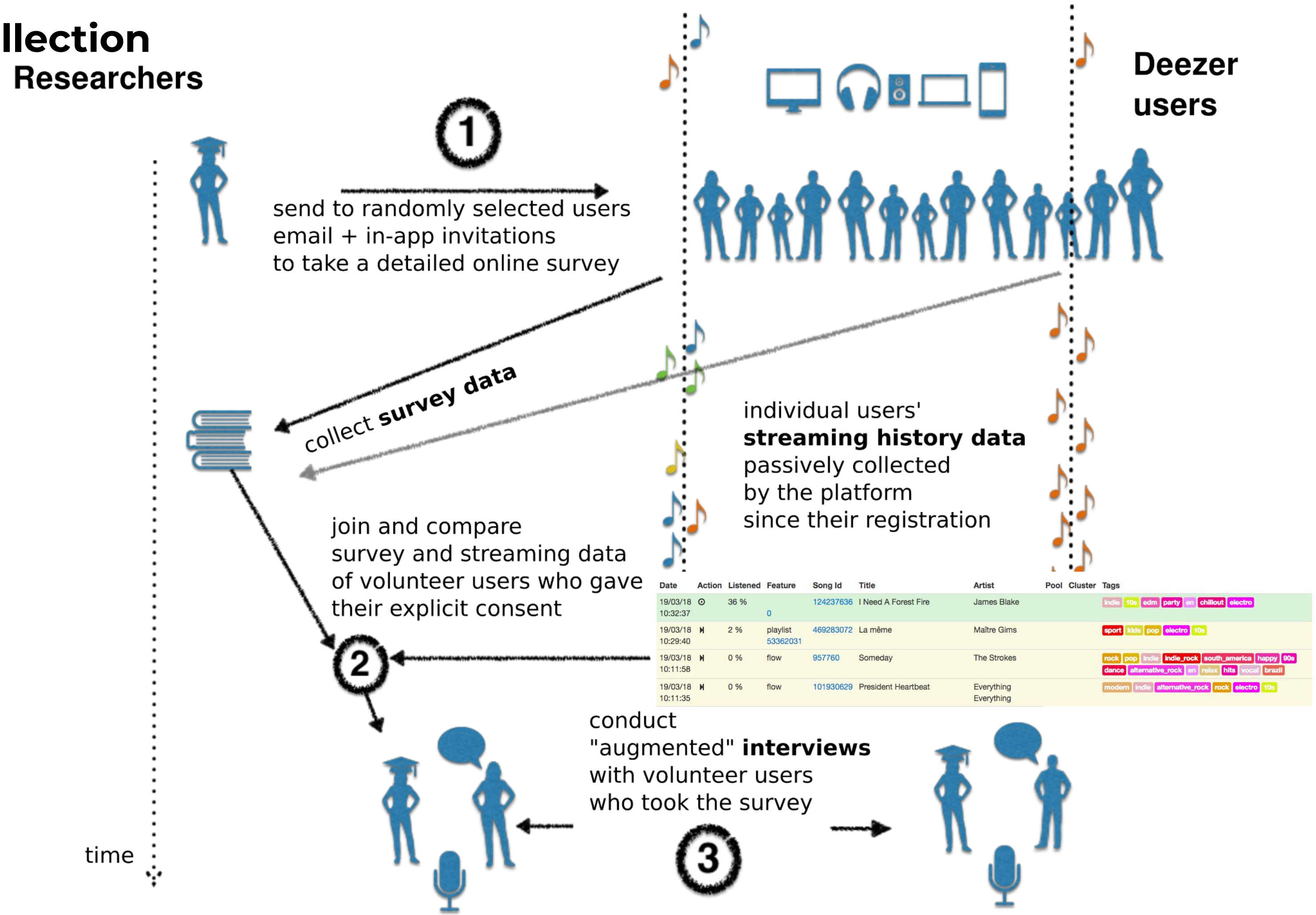
Sources: IFPI, Global music reports 2022, 2023
+ SNEP decoding (IFPI data)

First study //
Use cases of a mixed methods research design that integrates digital traces, survey data, and nested sampling

> Yann Renisio, Amélie Beaumont, Jean-Samuel Beuscart, Samuel Coavoux, Philippe Coulangeon, et al.. (2024) Integrating digital traces into mixed methods designs: An application to the study of online music listening using survey, interview and stream history data collected from the same people *Under review* ⟨hal-04448365⟩

RECORDS data collection

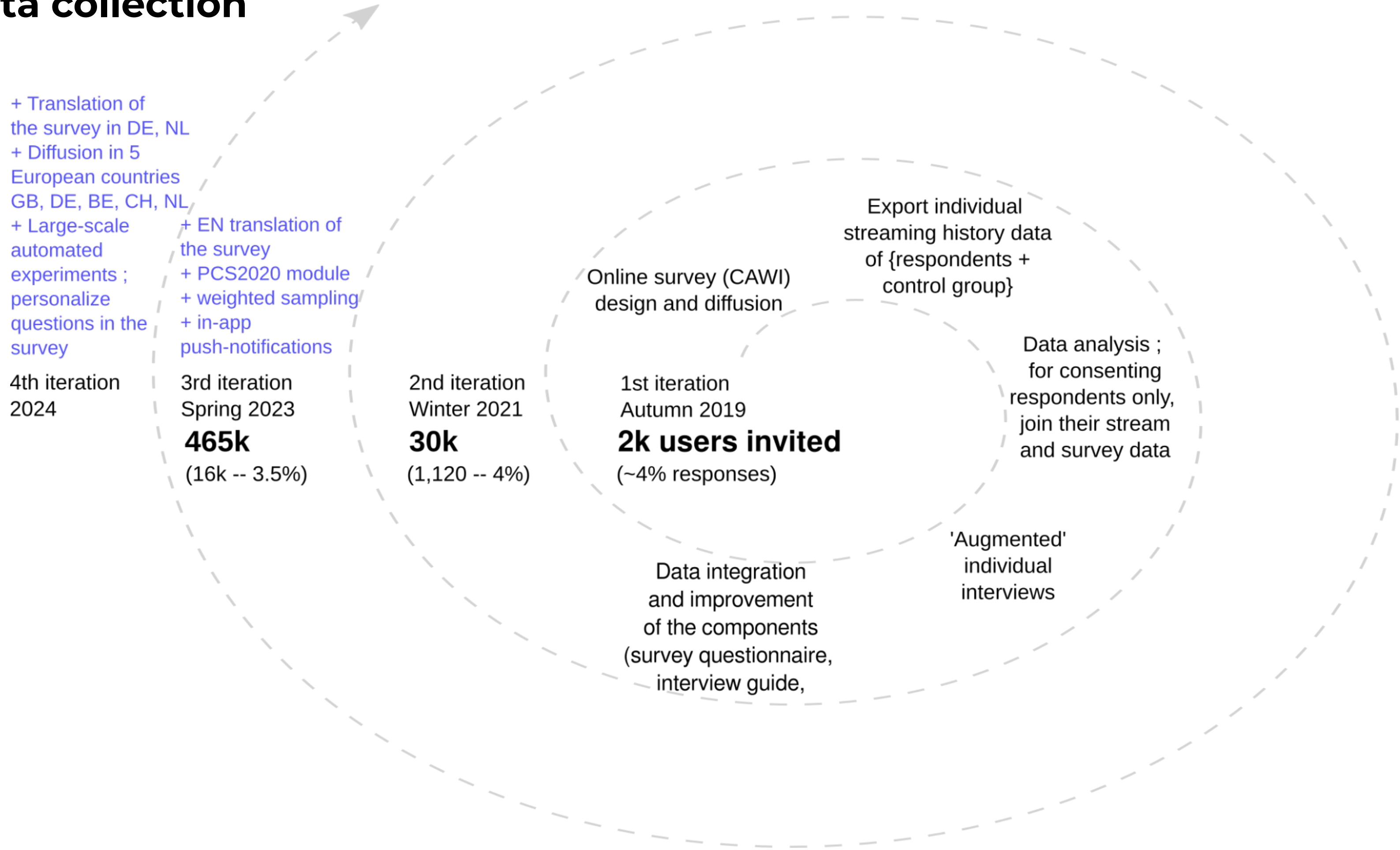
- **mixed**
- iterative
- nested



Date	Action	Listened	Feature	Song Id	Title	Artist	Pool	Cluster	Tags
19/03/18 10:32:37		36 %	0	124237636	I Need A Forest Fire	James Blake			indie 10s edm party en chillout electro
19/03/18 10:29:40		2 %	playlist 53362031	469283072	La même	Maître Gims			sport kids pop electro 10s
19/03/18 10:11:58		0 %	flow	957760	Someday	The Strokes			rock pop indie indie_rock south_america happy 90s dance alternative_rock en relax hits vocal brazil
19/03/18 10:11:35		0 %	flow	101930629	President Heartbeat	Everything Everything			modern indie alternative_rock rock electro 10s

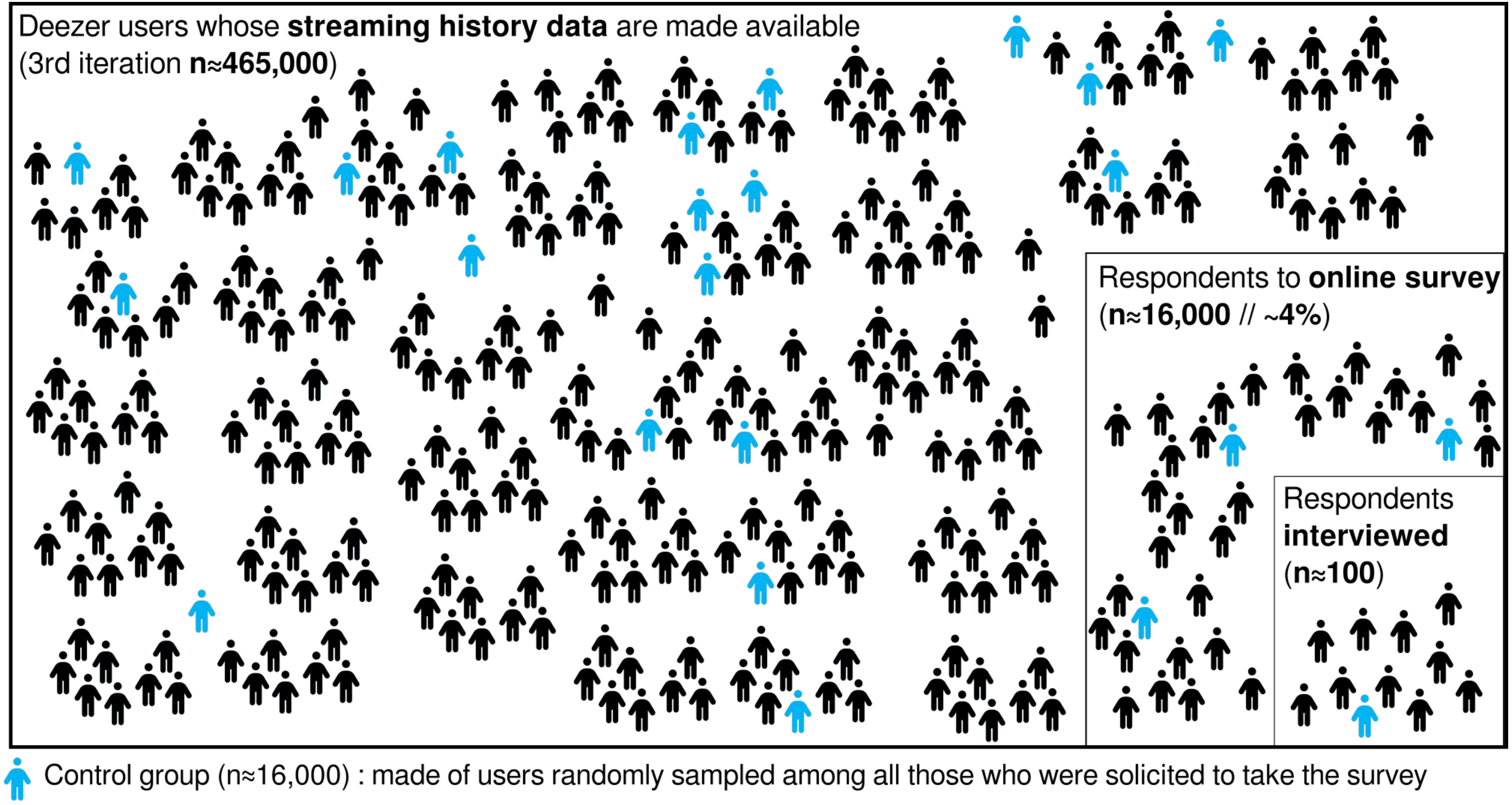
RECORDS data collection

- mixed
- **iterative**
- nested



RECORDS data collection

- mixed
- iterative
- **nested**



Use cases of the pairwise combination of data sources

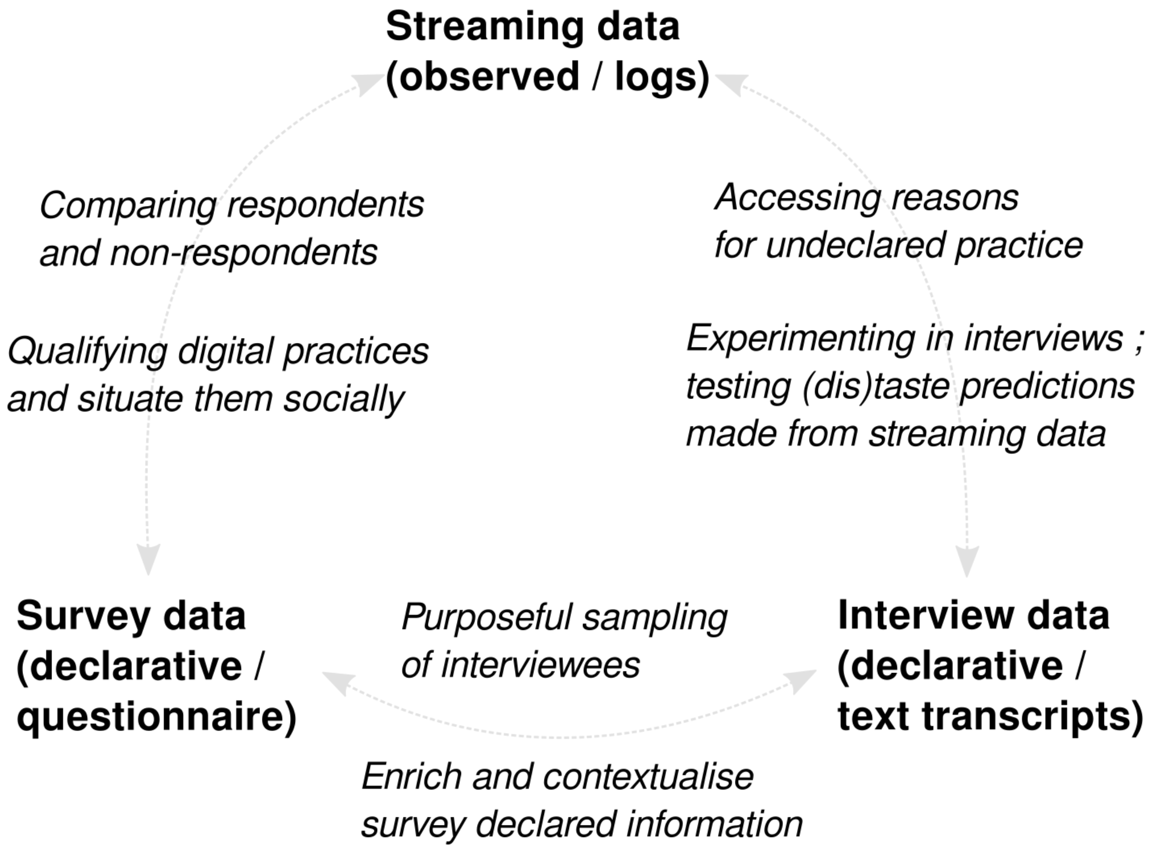
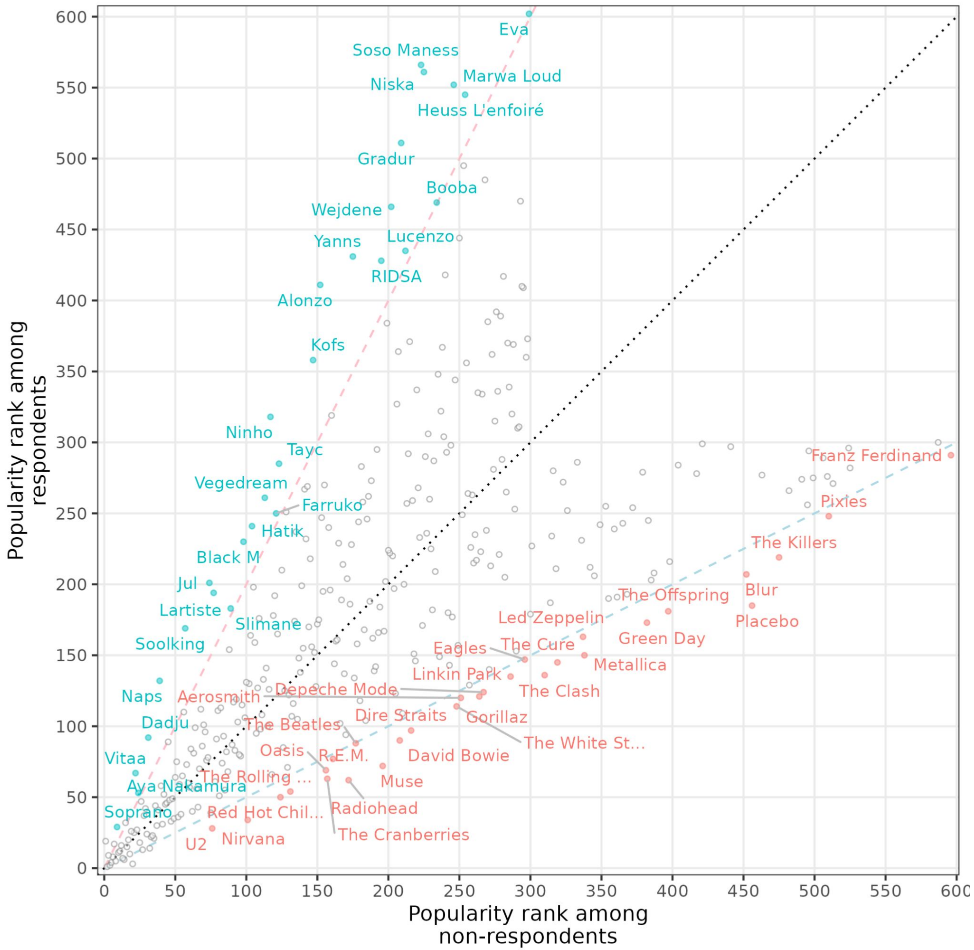


Illustration (1/2)



Streaming data (observed / logs)

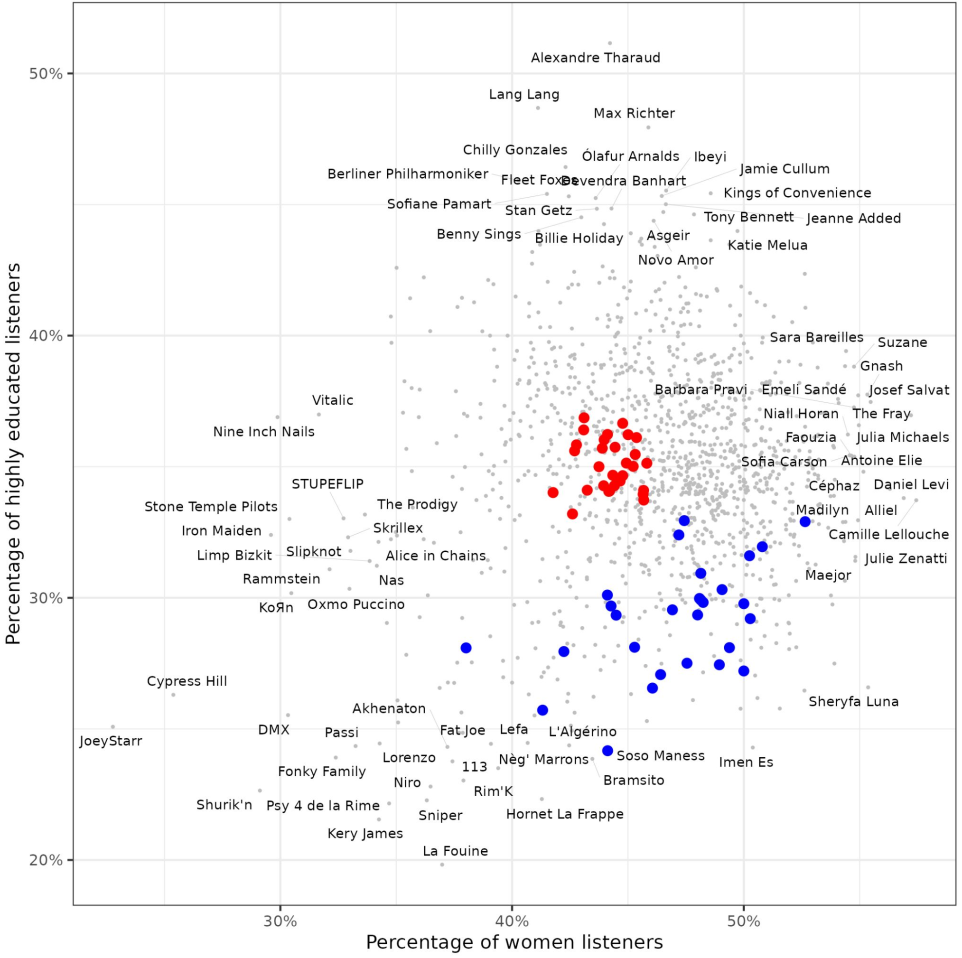
Comparing respondents and non-respondents

Qualifying digital practices and situate them socially

Survey data (declarative / questionnaire)

> Renisio et al.. (2024) Integrating digital traces into mixed methods designs: An application to the study of online music listening using survey, interview and stream history data collected from the same people. *Under review* ⟨hal-04448365⟩

Illustration (2/2)



Streaming data (observed / logs)

Comparing respondents and non-respondents

Qualifying digital practices and situate them socially

Survey data (declarative / questionnaire)

Purposeful sampling of interviewees

> Renisio et al.. (2024) Integrating digital traces into mixed methods designs: An application to the study of online music listening using survey, interview and stream history data collected from the same people. *Under review* ⟨hal-04448365⟩

Second study //
Unveiling social desirability scales

> Thomas Louail and Yann Renisio, Unveiling social desirability scales by comparing individuals’ responses to an online survey with their streaming history data. *In preparation*.

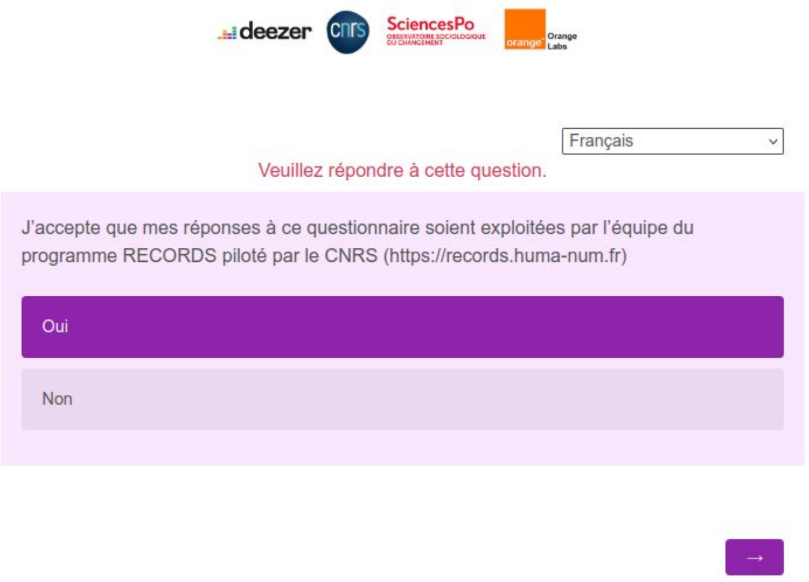
Motivation (1/2)
Build on the gap between practice and declaration of practice

- Digital traces are sometimes presented as able to replace survey data (on the long run), because considered a better proxy for actual practice than the declarations and self-reports collected in surveys often sought to capture (Parry et al. 2021)
- When the two sources are available and used together,
  it is often to 'improve'/enrich one with the other (Stier et al. 2019)
  - correct social desirability biases in survey data
  - reduce noise in observational digital traces
- Our approach: rather than 'triangulating', we want to use the measure of discrepancies to build new indicators to compare measurement categories
  - compare discrepancies for different categories of content
  - highlight social norms related to these categories
- Two well-known cases of social desirability biases:
  Body Mass Index; Voter abstention

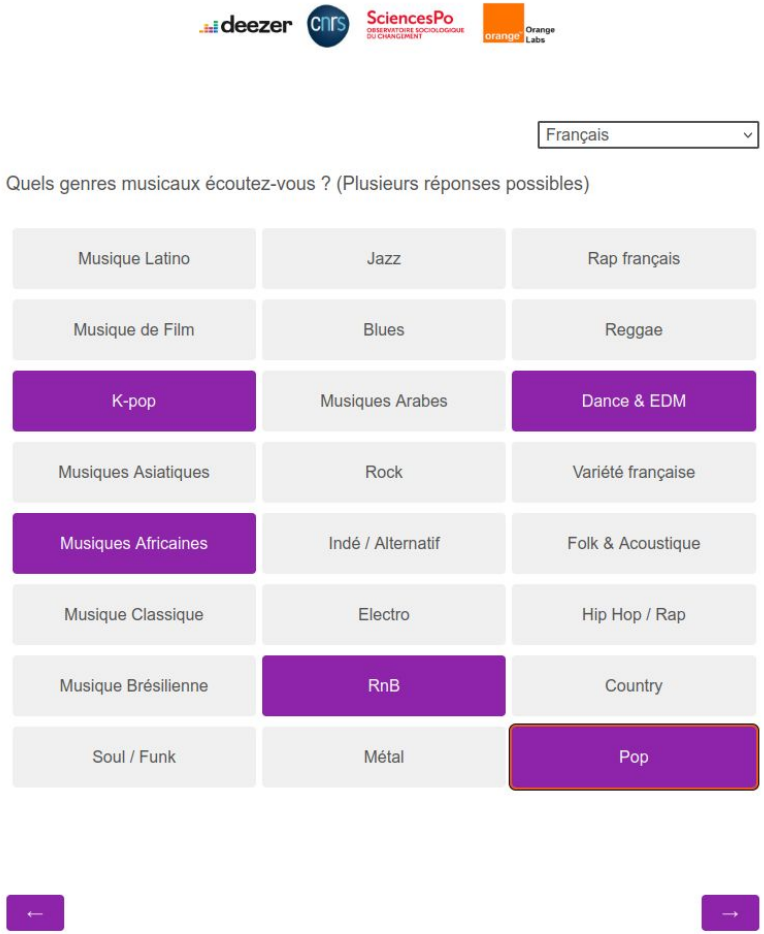
Motivation (2/2)
Application to music listening and preferences

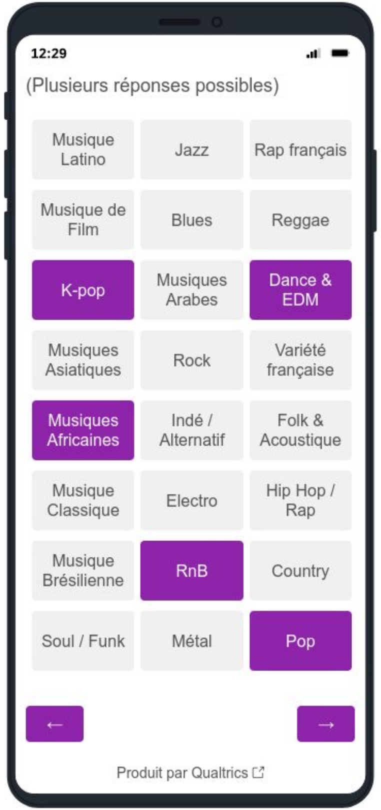
- The case of music consumption makes it possible to elaborate on the numerous studies that have highlighted the statistical affinity between the music preferences declared by respondents and their social positions, e.g. (Coulangeon 2003) for France ; (Peterson & Kern 1996) for the US ; (Chan & Goldthorpe 2007) for the UK
- In particular it has been long documented that:
    - *Classical music* is more often declared to be listened to by the upper classes;
    - *Rap*, on the contrary, is more often listened to by the working classes;
    - *Jazz* and *Metal* are somewhat intermediary between these two genres: C > J > M > R
- Would we retrieve this order if we compared the genres according to how much they are over-/under-declared when compared to how much they are streamed?

Survey snapshots (1/3) :
Asking users' explicit consent

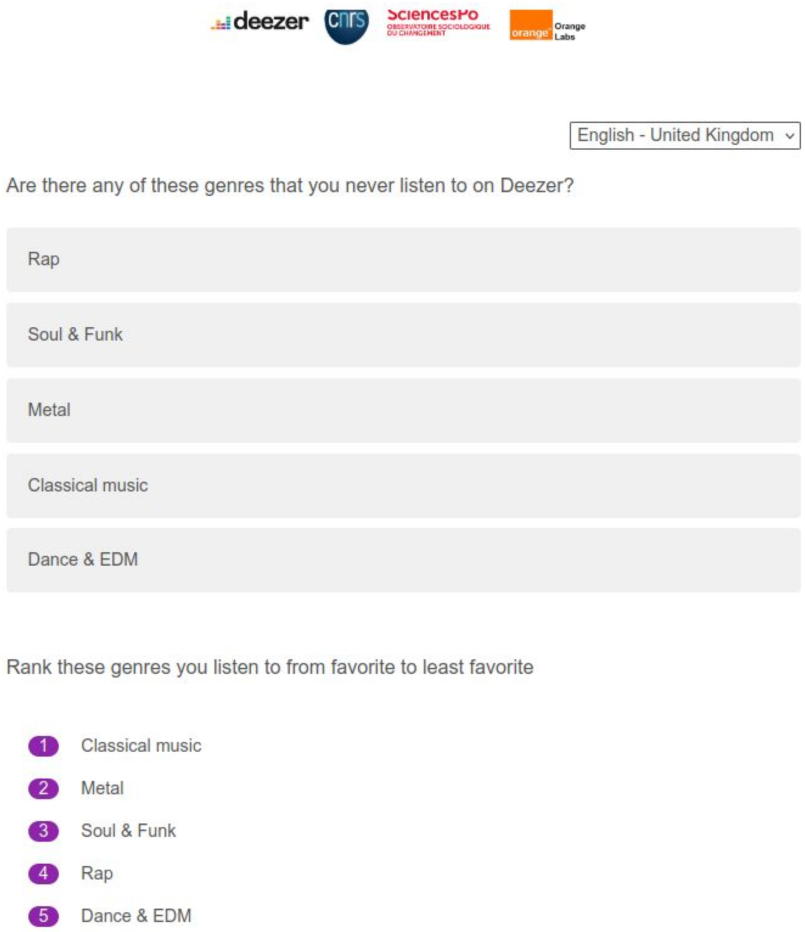


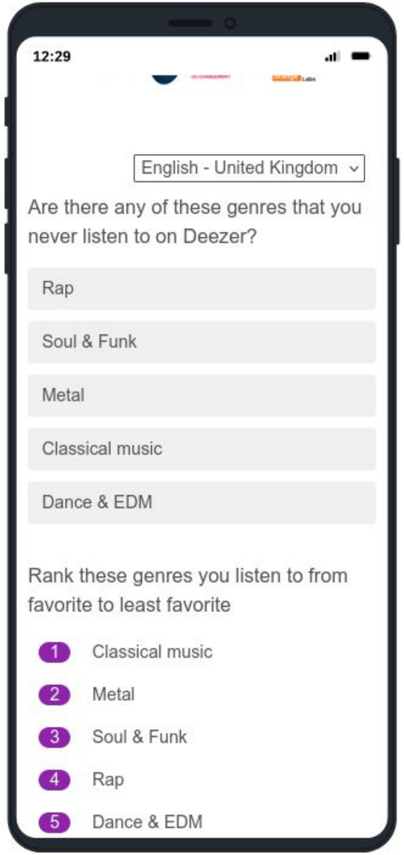
Survey snapshots (2/3) :
Asking about listening habits and music genres





Survey snapshots (3/3) :
Asking about listening habits and music genres

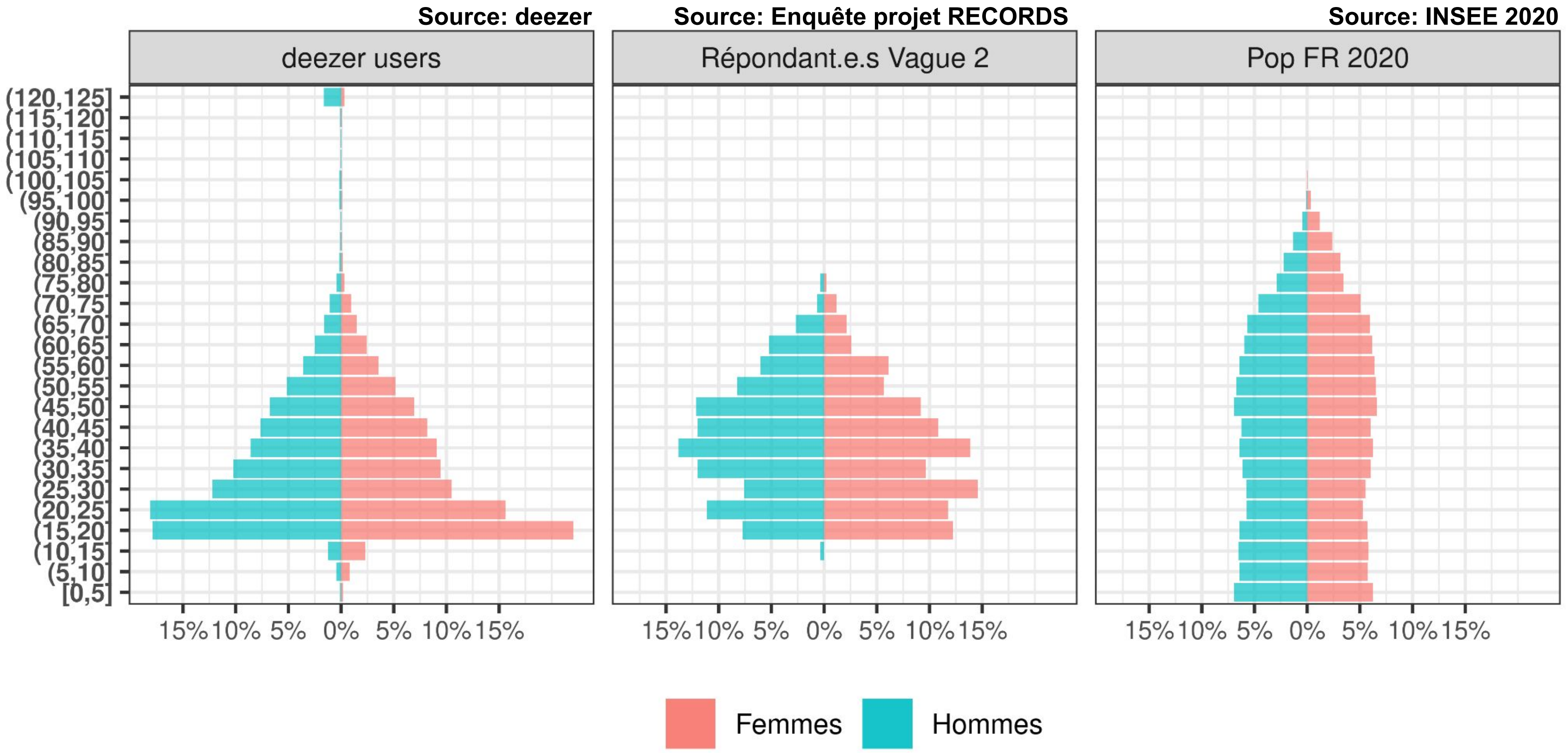




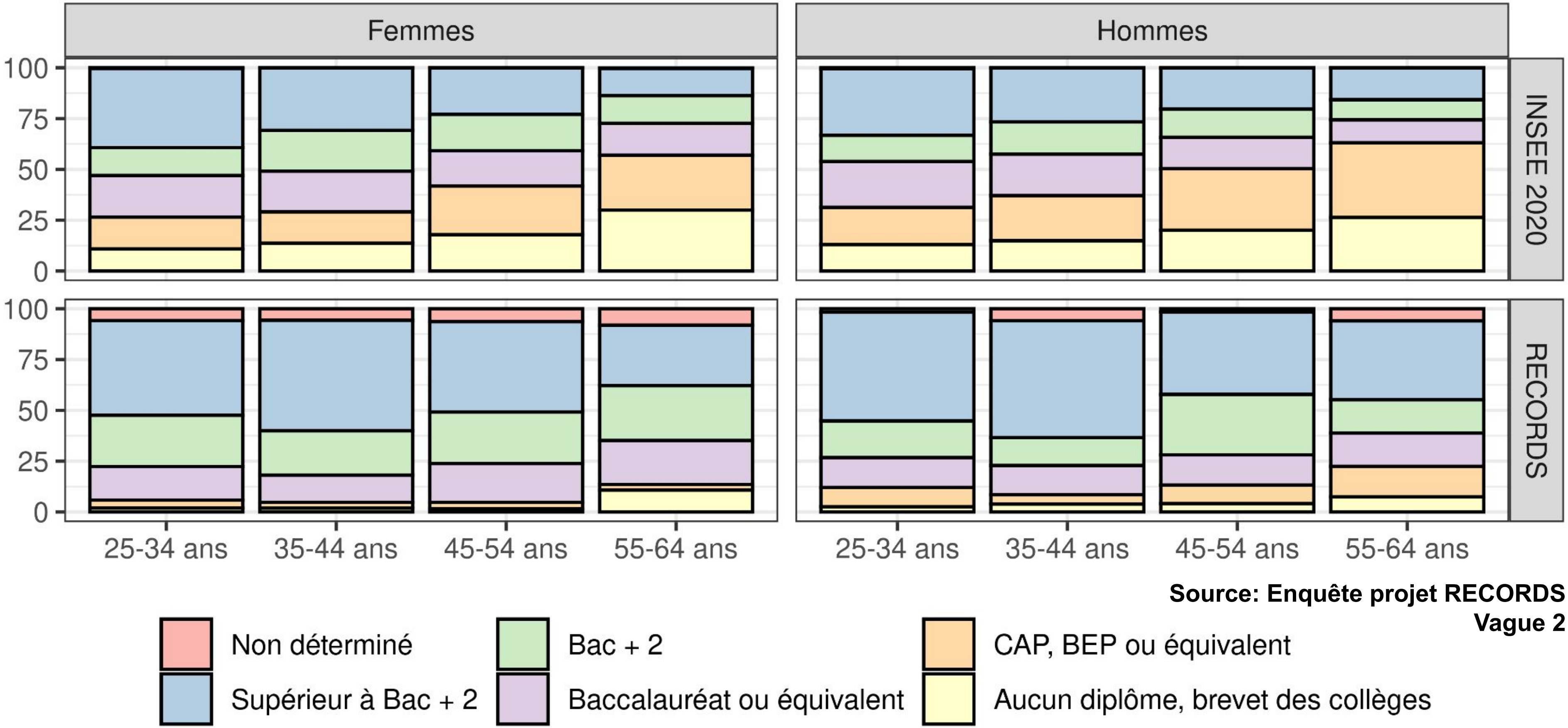
The population of survey respondents

Iteration n°2 // January 2021
30k users of Deezer located in France
email invitation to take the survey
no incentive
N=1.4k respondents (~4%)

Demographics of a (bad) sample



Highest qualification level (highest diploma obtained)



Source: Enquête projet RECORDS Vague 2

Methods

1. Assign musical genres to the millions of tracks and artists streamed by the survey respondents
2. Systematically compare the weight of these genres in the individuals' streams with the listening habits they reported in the survey, and rank
3. Rank musical genres based on the intra-individual deviations between reported listening habits and actual streams
4. Apply the same method to artist preferences

Methods

1. Assign musical genres to the millions of tracks and artists streamed by the survey respondents
2. Systematically compare the weight of these genres in the individuals' streams with the listening habits they reported in the survey
3. Rank musical genres based on the intra-individual deviations between reported listening habits and actual streams
4. Apply the same method to artist preferences

Tracks, artists and genres (1/2)

People listen to songs/music pieces and artists, whose qualification by one or more genre labels (rock, jazz, rap, etc.) is debatable

> Genres are problematic and ill-defined categories
* no consensus on a definition (Robette & Roueff 2014)
* they aggregate very different products (Nault et al. 2021)
* their boundaries evolve over time (Levine, 1988)

=> more formally, there’s no unique and well-defined function
$F: t \rightarrow g$
that would associate a finite number of genre labels g
to each track (or artist) t

BUT almost the entire scientific literature in sociology of music preferences is written at the level of music genres…

Highbrow: classical and jazz
Middlebrow: rock
Lowbrow: dance and rap

→ There exists a very large number of reasonable choices for the labeling function F

Tracks, artists and genres (2/2)

Over the past 12 months prior to the survey, for the 16k respondents to the 3rd iteration of the survey :

313k distinct artists streamed at least once

3M unique tracks

=> Manual labeling is out of reach

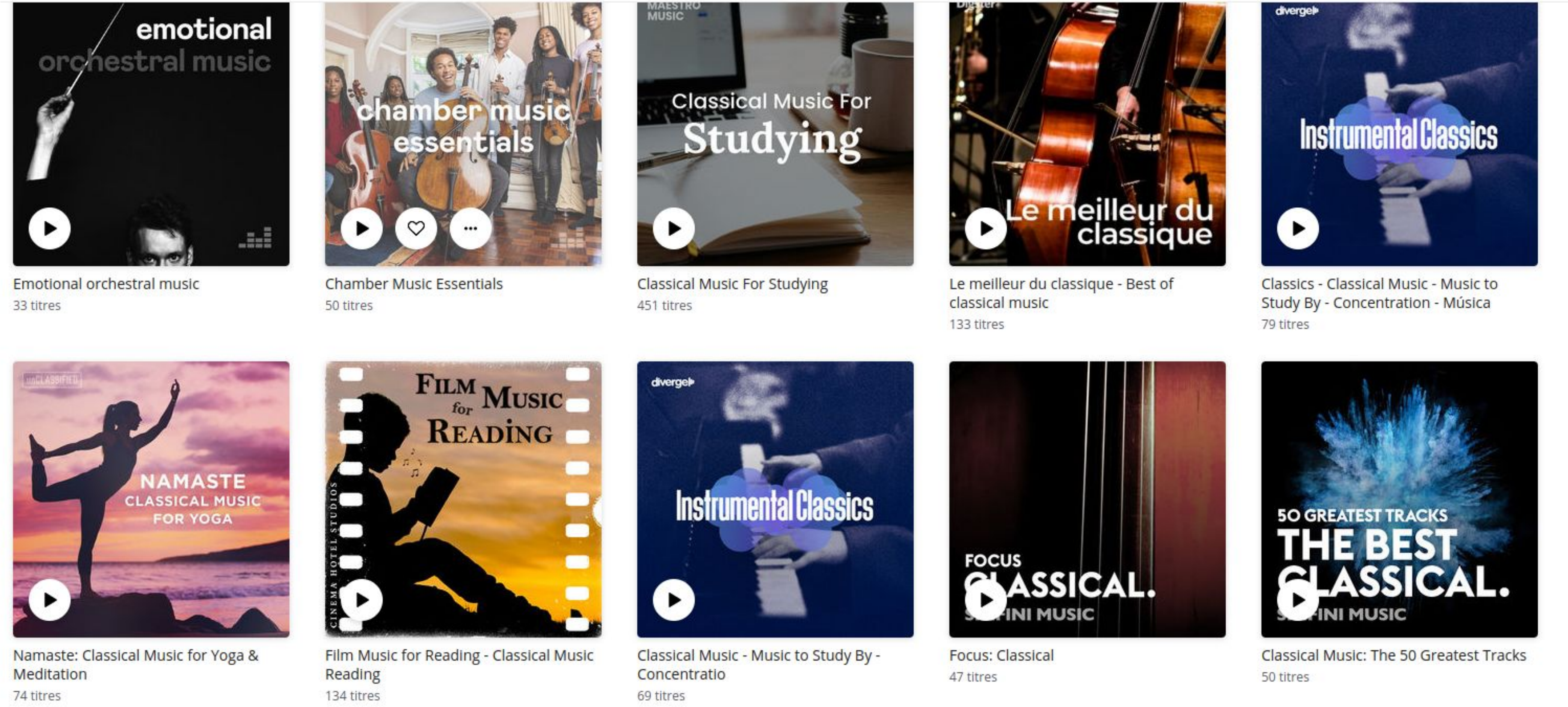
Different systematic approaches to labeling:
- from online collaborative databases (e.g. musicbrainz, wikidata)
- from the content of user playlists named after a music genre
- from the content of editorial playlists named after a music genre
- from the programming of French radio stations dedicated to a given music genre (France Musiques, Radio Classique, TSF Jazz, etc.)
- ...

In the following, to simplify we reduce a music genre to a set of artists whose tracks are often associated with this genre in playlists curated by professional editors (employed by the streaming platform)

Assumption: a strong result on possible deviations between genres, in how people declare and/or listen to them, should hold whatever the procedure chosen for assigning genre labels to musical items

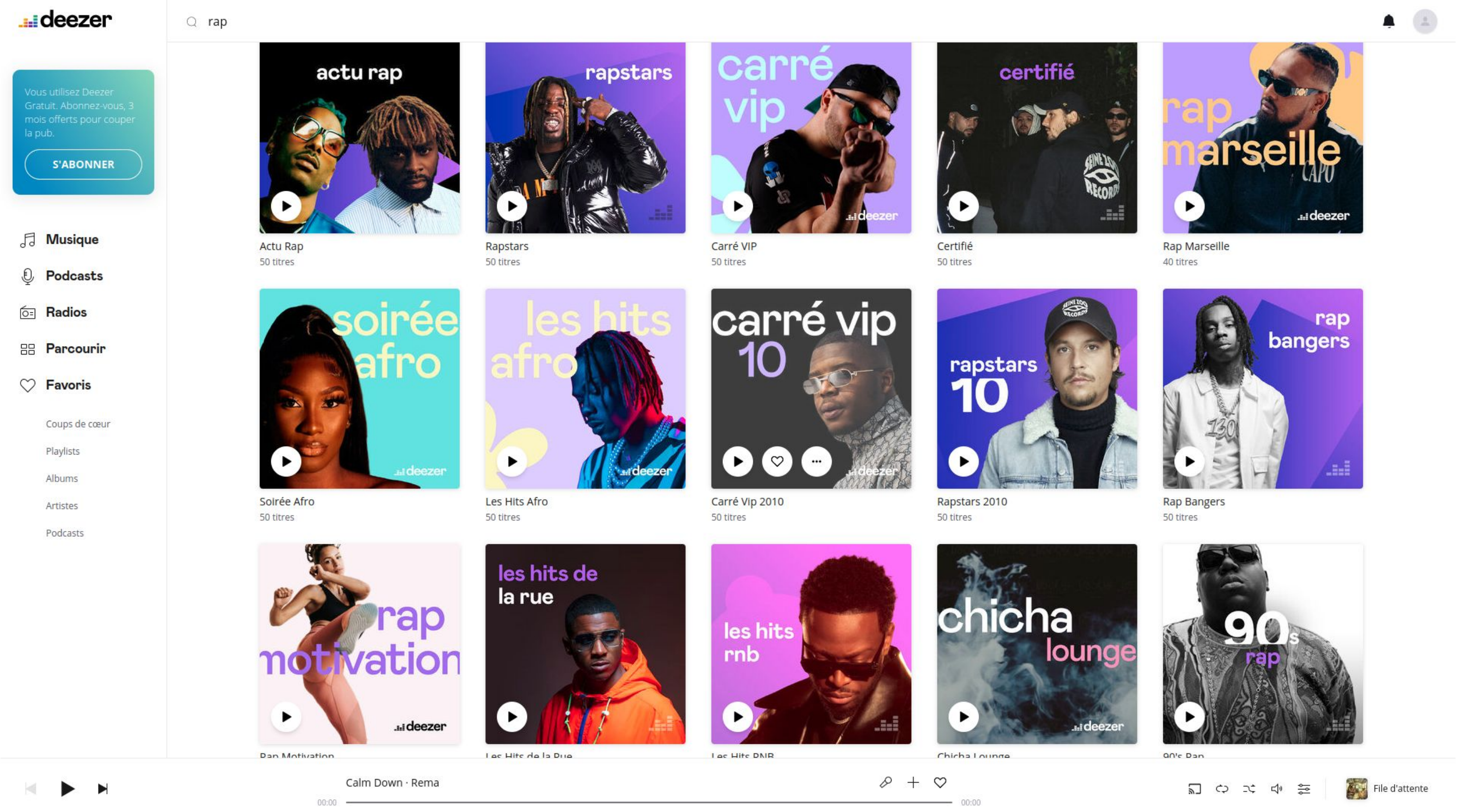
An expert-based and statistical approach

We rely upon the **editorial playlists** proposed by the platform that are named after a music genre



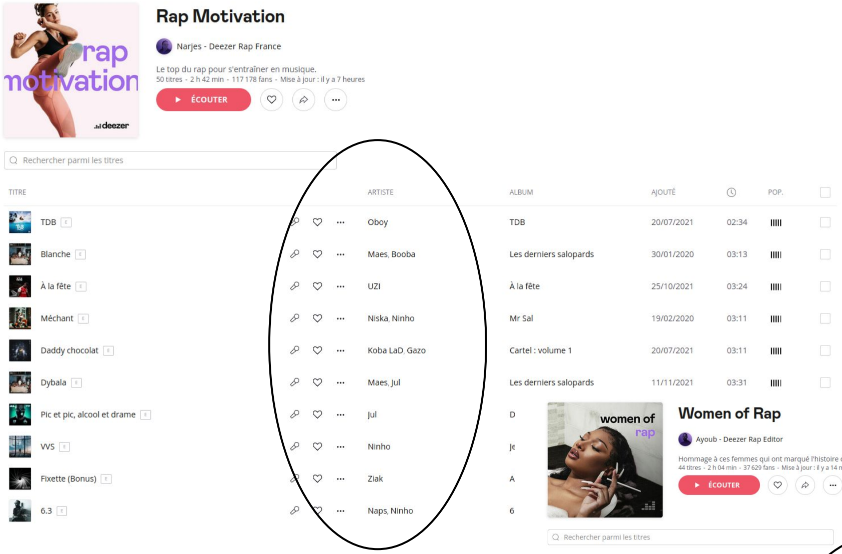
An expert-based and statistical approach

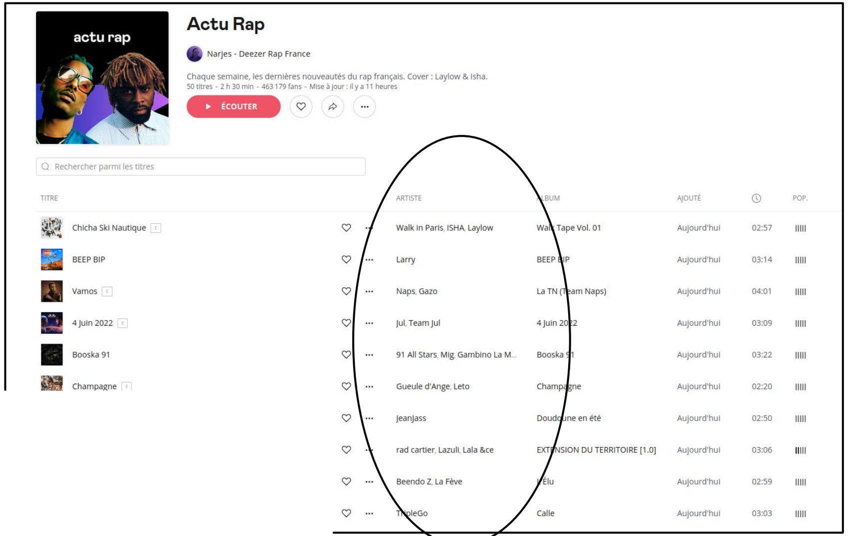
We rely upon the **editorial playlists** proposed by the platform that are named after a music genre

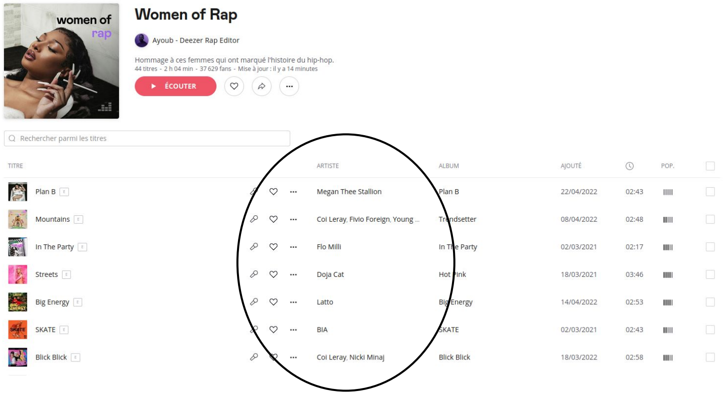


An expert-based and statistical approach

We rely upon the **editorial playlists** proposed by the platform that are named after a music genre







Consistency of the lists of artists per genre

- 80% of artists are tagged, representing ~91% of respondents' streams in the last 12 months
- The most frequently featured artists in these editorial playlists named after a genre are easily identifiable as popular artists of that genre →

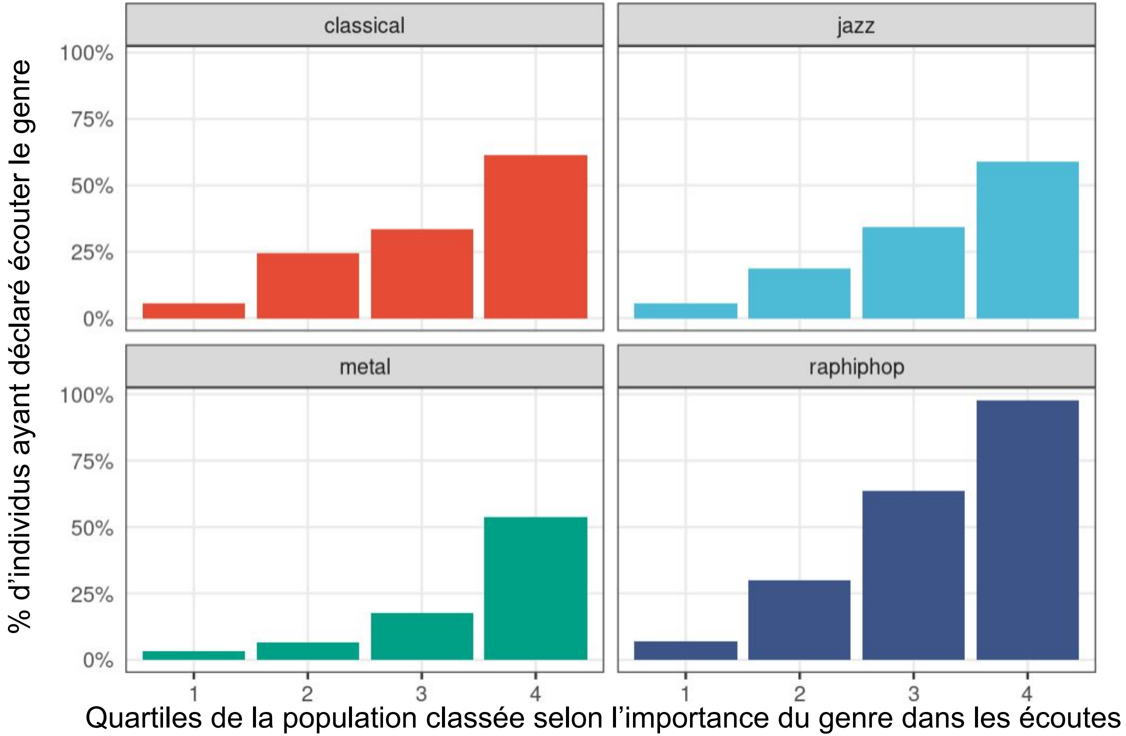
rank	classical	jazz	metal	raphiphop
1	Lang Lang	Ella Fitzgerald	Metallica	Drake
2	Berliner Philharmoniker	Miles Davis	Iron Maiden	Eminem
3	Daniel Barenboim	Louis Armstrong	Megadeth	Kanye West
4	London Symphony Orchestra	Nina Simone	Slipknot	Travis Scott
5	Max Richter	John Coltrane	Judas Priest	Post Malone
6	Ludovico Einaudi	Stan Getz	Black Sabbath	Dr. Dre
7	Wiener Philharmoniker	Herbie Hancock	Slayer	Snoop Dogg
8	KHATIA BUNIATISHVILI	Diana Krall	Pantera	Future
9	Martha Argerich	Billie Holiday	Sepultura	50 Cent
10	Alexandre Tharaud	Frank Sinatra	KoЯn	Lil Wayne
11	Yann Tiersen	Chet Baker	Motörhead	Migos
12	Yo-Yo Ma	Norah Jones	Disturbed	Kendrick Lamar
13	John Williams	Bill Evans	Ozzy Osbourne	Jay-Z
14	Joep Beving	Duke Ellington	System of a Down	Nas
15	Riccardo Muti	Nat King Cole	Accept	DJ Khaled
16	Alice Sara Ott	Sarah Vaughan	Helloween	J. Cole
17	Anne-Sophie Mutter	Gregory Porter	Arch Enemy	2Pac
18	Dustin O'Halloran	Ray Charles	Anthrax	Booba
19	Hans Zimmer	Tony Bennett	Nightwish	XXXTentacion
20	Hélène Grimaud	Sonny Rollins	Scorpions	Chris Brown

Methods

1. Assign musical genres to the millions of tracks and artists streamed by the survey respondents
2. Systematically compare the weight of these genres in the individuals' streams with the listening habits they reported in the survey
3. Rank musical genres based on the intra-individual deviations between reported listening habits and actual streams
4. Apply the same method to artist preferences

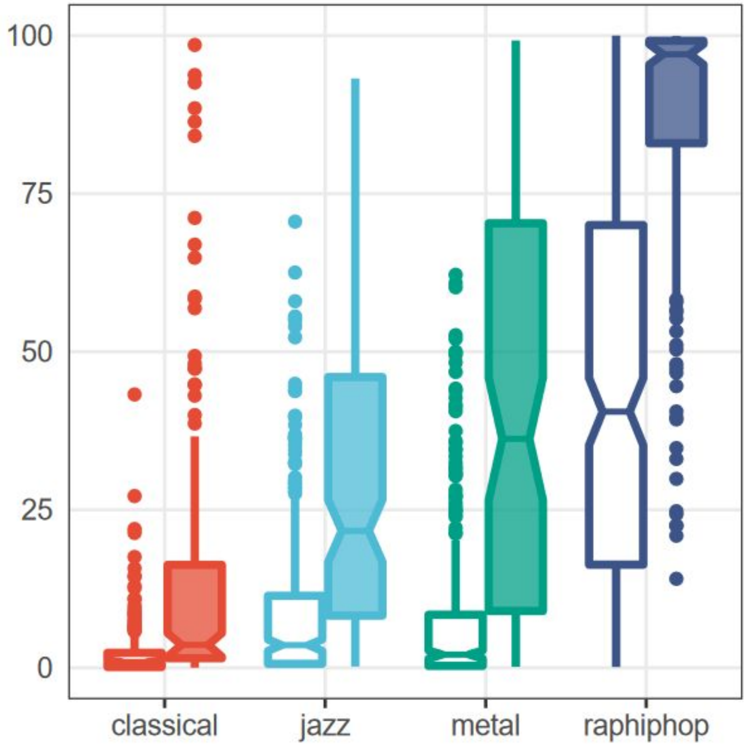
The more respondents stream a given music genre, the more they declare to listen to that genre

- This holds whatever the music genre considered



And individuals who claim to listen to a genre listen to it more than those who claim not to listen to it

- It is true regardless of the music genre



Measuring intra-individual discrepancies

We define in a non-ambiguous way an inconsistent statement:

If individual i declares listening to genre G and not genre G',
there is inconsistency if this individual listens to more G' than G
Additionally, the intensity of this difference is obtained by the ratio of
the times he/she spent streaming the two genres: t(G)/t(G')

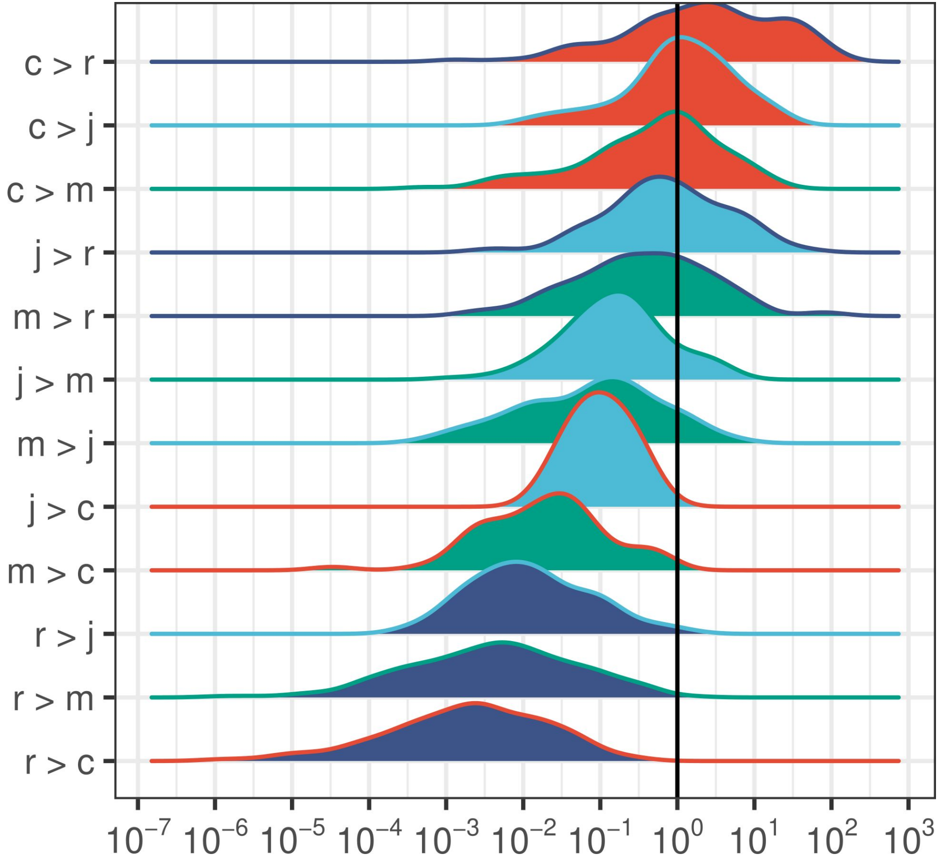
For example, an individual who claims not to listen to classical music
and to listen to rap will be considered inconsistent if he listens to more classical music than rap.
If he has listened to 3 hours of classical music and 1 hour of rap, the intensity of this difference will be 3/1

The unequal social desirability of musical genres

However, we observe rather the opposite: the majority of respondents who declare listening to classical music and not rap, listen to less classical music than rap

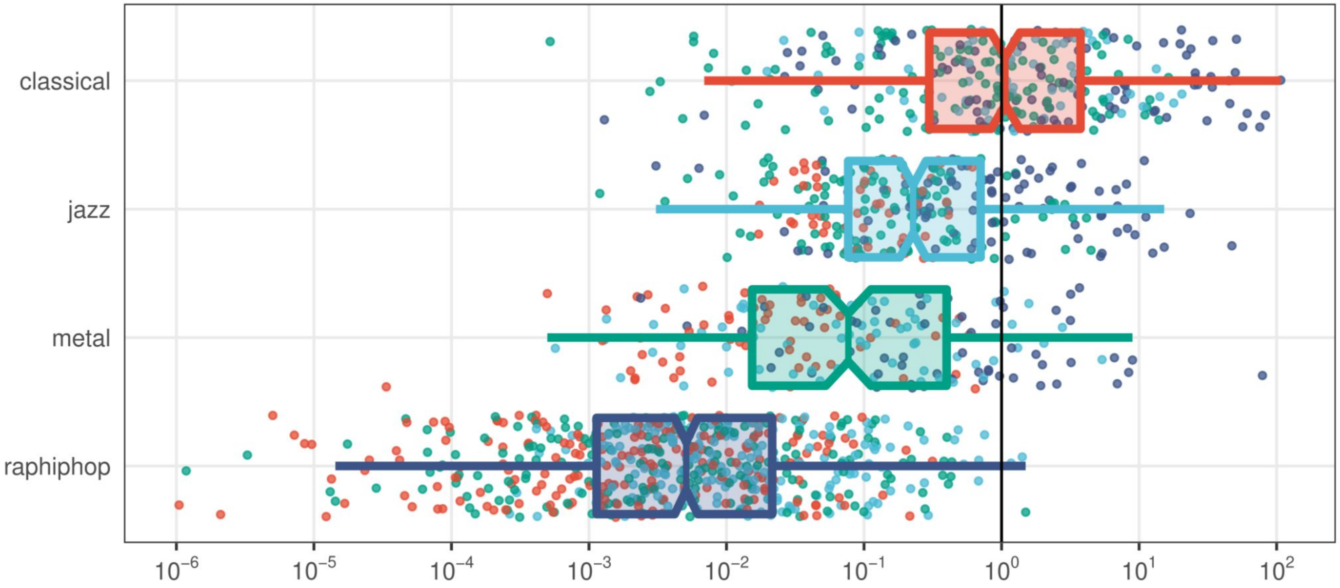
Conversely, almost all individuals declaring to listen to rap and not classical music have a streaming practice that align with their declaration

A gradation can thus be observed between the pairs of genres.



The unequal social desirability of musical genres

By aggregating these results, we retrieve a ranking of genres often discussed in the literature, but based solely on the discrepancies between survey responses and digital traces (rather than on the social properties of respondents).



Methods

1. Assign musical genres to the millions of tracks and artists streamed by the survey respondents
2. Systematically compare the weight of these genres in the individuals' streams with the listening habits they reported in the survey
3. Rank musical genres based on the intra-individual deviations between reported listening habits and actual streams
4. Apply the same method to artist preferences

Excerpt of the online survey

B4. Quels genres musicaux écoutez-vous ? (Plusieurs réponses

- [ ] Rap français
- [x] Pop
- [x] Rock
- [ ] Electro
- [ ] Jazz
- [ ] RnB
- [x] Soul / Funk
- [x] Hip Hop / Rap
- [x] Variété française
- [ ] Musique Classique
- [ ] Reggae
- [ ] Métal
- [ ] Indé / Altern
- [ ] Musique de
- [ ] Folk & Acou
- [ ] Musique Lat

B5-variete. Que pensez-vous des artistes de variété française suivants ?

	je déteste	je n'aime pas	je ne connais pas	pas d'avis	j'aime	j'aime beaucoup
Georges Brassens	○	○	○	○	●	○
Jean-Jacques Goldman	○	○	○	○	●	○
Michel Sardou	○	○	○	●	○	○
Renaud	○	○	○	○	●	○
Calogero	○	●	○	○	○	○
Céline Dion	○	●	○	○	○	○
Johnny Hallyday	○	○	○	●	○	○

B5-rock. Que pensez-vous des artistes de Rock suivants ?

	je déteste	je n'aime pas	je ne connais pas	pas d'avis	j'aime	j'aime beaucoup
U2	○	○	○	○	●	○
Metallica	○	●	○	○	○	○
Imagine Dragons	○	○	○	○	○	●
Oasis	○	○	○	○	●	○
The Beatles	○	○	○	○	●	○
Led Zeppelin	○	○	○	○	●	○
Luke Carlson	○	○	●	○	○	○
Radiohead	○	○	○	○	○	●
Queen	○	○	○	○	●	○
AC/DC	○	○	○	○	○	●
Twenty One Pilots	○	○	○	○	○	●
The Clash	○	○	●	○	○	○
PJ Harvey	○	○	●	○	○	○
Téléphone	○	○	○	○	●	○

Artist overstatements and social desirability

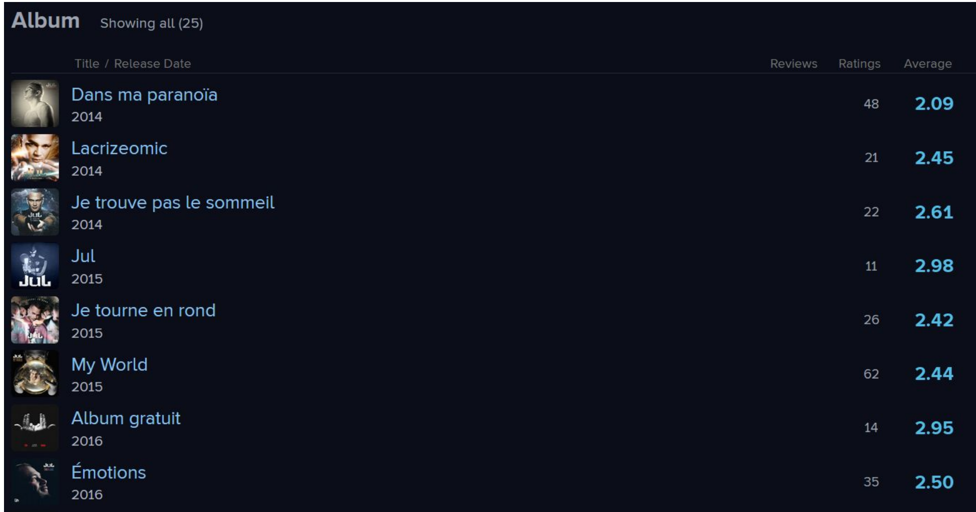
To extend this principle on better defined ‘objects’, we tested the same hypothesis on artists. For each respondent, and for each pair of artists in the survey questionnaire, we hypothesized that if the respondent (implicitly) declared preferring A to B, he/she should listen to A more than B.
We then aggregate the results by artist to calculate an artist's overstatement rate.

In parallel, for each artist in the questionnaire, we extract from rateyourmusic.com the ratings given by the website’s contributors to the artist's albums, and calculate the average rating (weighted by the number of reviews) as a proxy for recognition.

!! Number of reviews varies greatly from one artist to another

Que pensez-vous des artistes de variété française suivants ?

	je n'aime pas	pas d'avis	j'aime
Céline Dion	○	●	○
Indochine	●	○	○
Serge Gainsbourg	○	○	●
Julien Doré	○	●	○
Alain Bashung	○	○	●
Georges Brassens	○	○	●
Henri Salvador	○	○	●
Jacques Brel	○	○	●
Christophe Maé	●	○	○



Album Showing all (25)

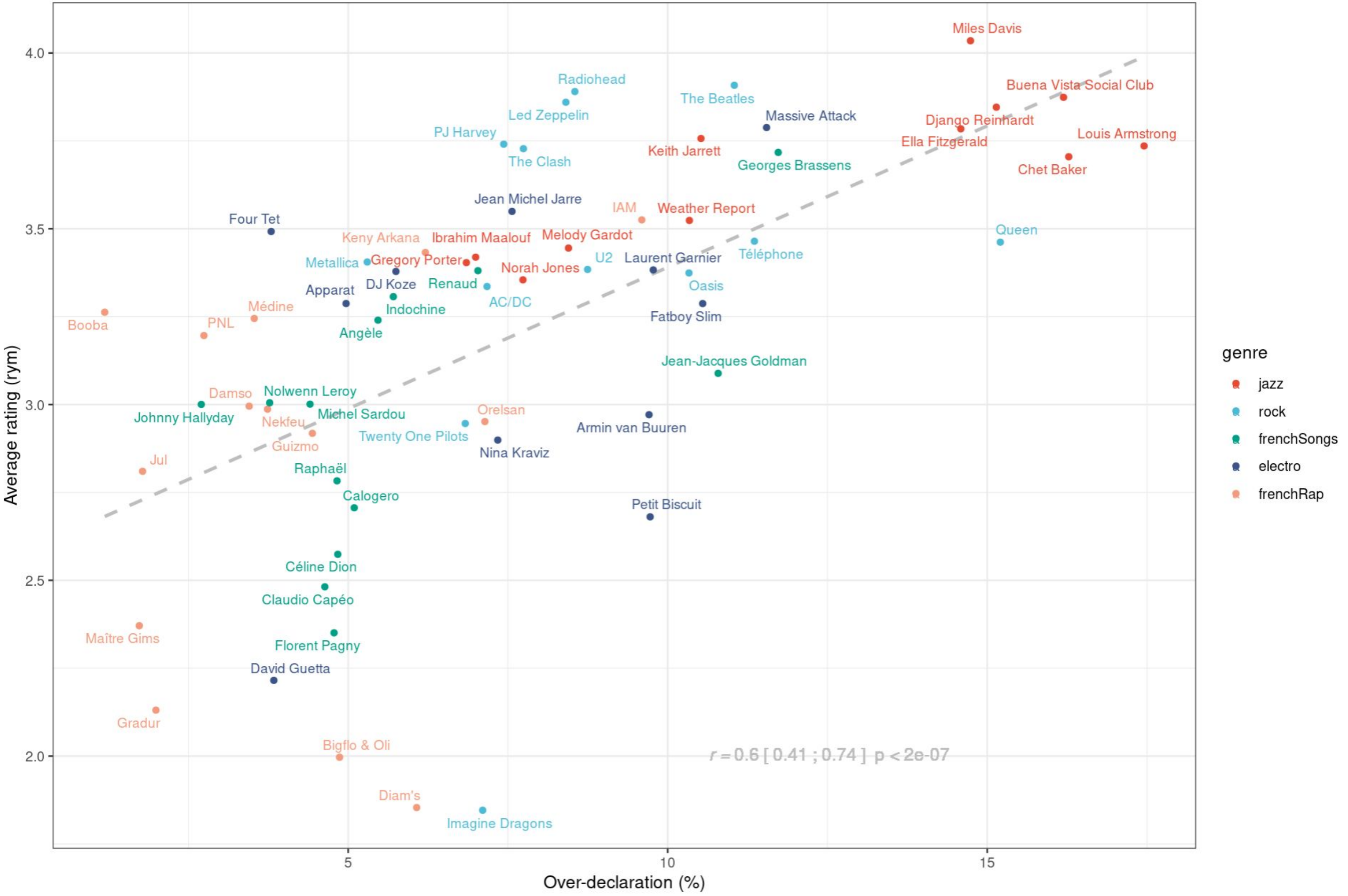
Title / Release Date	Reviews	Ratings	Average
Dans ma paranoïa 2014		48	2.09
Lacrizeomic 2014		21	2.45
Je trouve pas le sommeil 2014		22	2.61
Jul 2015		11	2.98
Je tourne en rond 2015		26	2.42
My World 2015		62	2.44
Album gratuit 2016		14	2.95
Émotions 2016		35	2.50

Hypothesis

The more an artist has good reviews, the more likely she/he will be declared as more appreciated than another artist, yet listened to more.

Overdeclarations of artists and social desirability

Some fluctuations
but a trend:
the artists most often over-declared as appreciated (relative to how much they are listened to) are those whose albums are the best rated



Still needs to be done

- Reproduce analyses using other genre labeling functions
- Does the propensity to conform to the norm vary significantly
  between social groups (gender, generation, socio-professional category, etc.)?
- Use different rating databases given to albums and (aggregated to artists)
  by both specialized media and internet users
- Reproduce calculations on RECORDS' 3rd iteration data
  - more volume (16k respondents)
  - better encoding of the socioeconomic status and educational attainment
  - filter listeners who declare listening to a music genre exclusively outside of Deezer
    (often heard: "people do not stream classical music, they listen to it on CDs with hifi
    material")
  - additional questions relative to genre appreciation and ordering
- Reproduce the analysis for users/respondents from different countries
  where Deezer is a big player in the market (Germany, UK and other EU countries)

Summary and outlook

- First exploration confirm the **value of working on the discrepancies between survey responses and digital traces to characterize social desirability scales**
- Research protocol requires collaboration with an internet platform willing to survey its users
  + when these users give their consent, willing to provide access to their individual digital traces
- Compared to other approaches — API ; data donation ; have users install apps that log their phone usage and web browsing – cf. F. Keusch (2024) Quality of Digital Behavioral Data. *FORS-SSP Methods & Research Meeting*
    - behavioral data recorded in natural conditions
    - very long time periods (several years for users who registered to the service long ago)
- The approach could be replicated to measure other kinds of digital consumption in streaming
  (e.g. movies and TV shows), and more generally any ordinary practice interfaced by digital platforms
  (eg. online shopping, book medical appointments, online dating, news consumption)
- Work needed to identify a proper level of aggregation of the personal data prior to their release,
  to guarantee privacy/anonymity of respondents
- Potential collaboration with FORS on the upcoming 4th iteration of the data collection
    - improve survey and data quality
    - experiments in survey methodology

Thank you!

Thomas Louail
CNRS Humanities and Social Sciences
Géographie-cités and PACTE (associate) research units
thomas.louail@cnrs.fr