

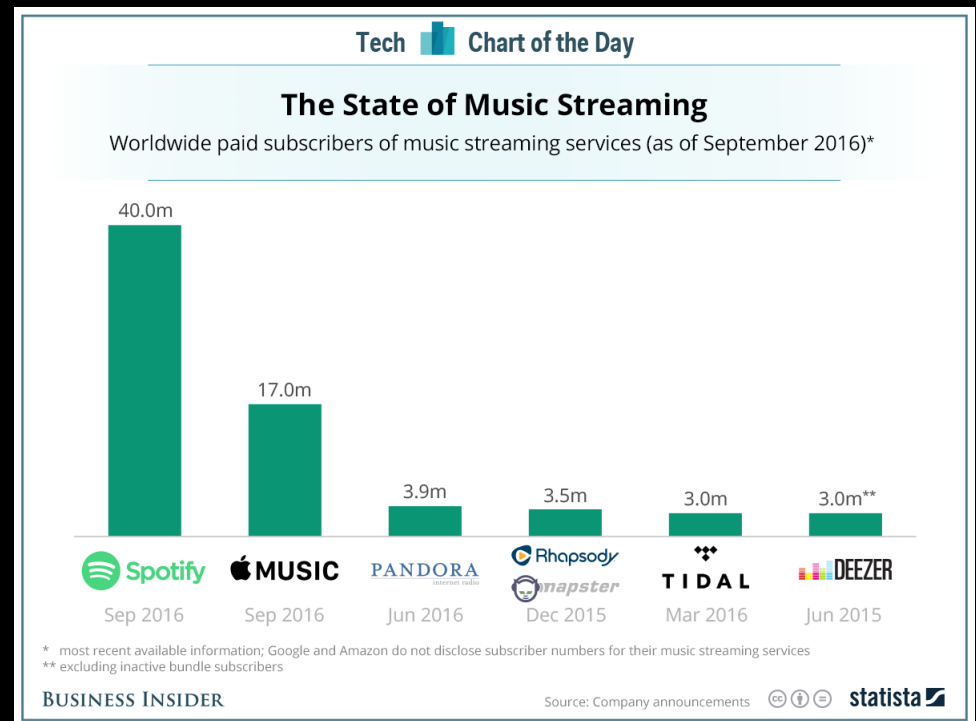
Comparison Top100 ranks from various service provider

Donghyun Kang

2018.02.13

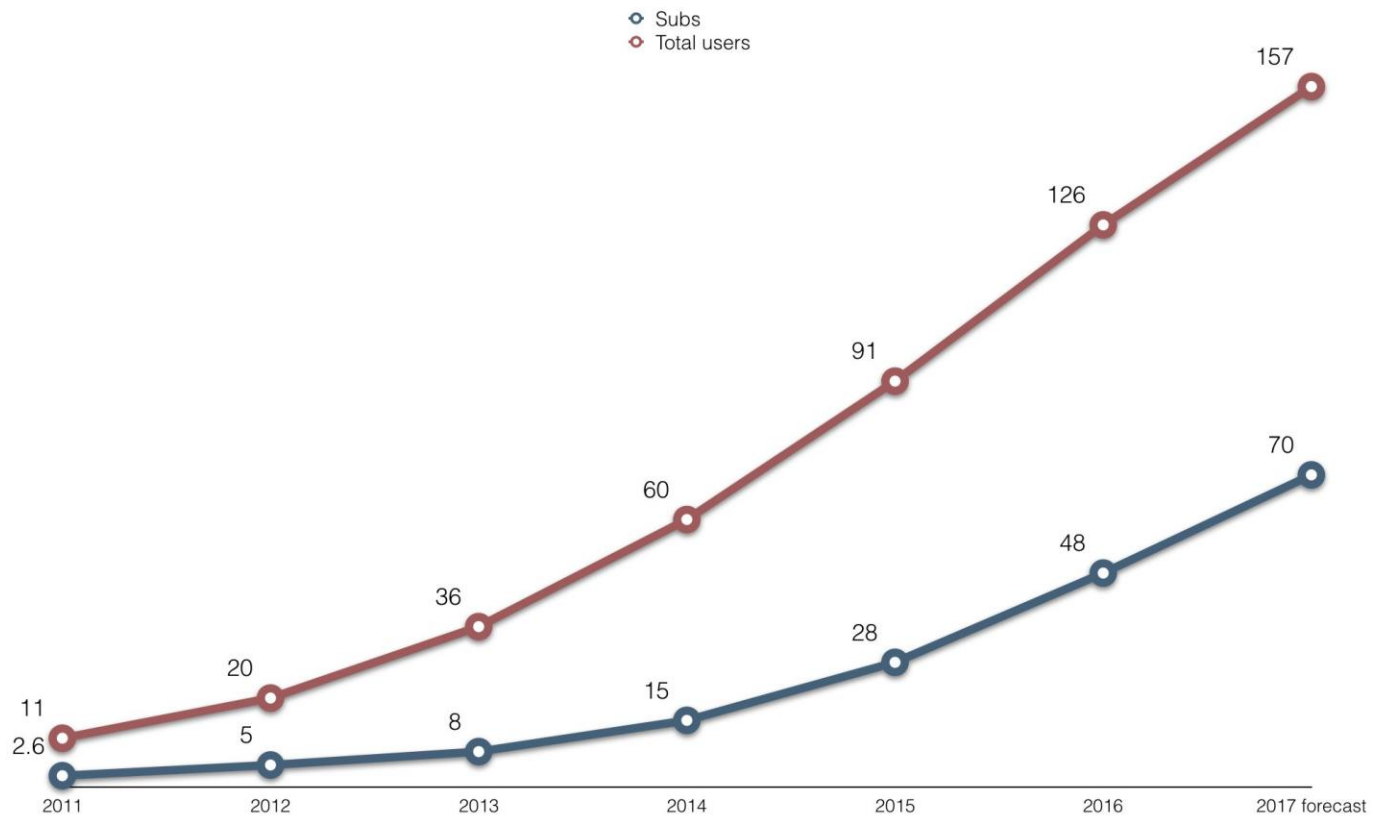
Data Scientist Fellow

NYC Data Science Academy



Motivation

Spotify end-of-year paying subs and total active users (m)



Source: Company accounts. 2014/2015 restated in 2016 accounts. 2017 is MBW's forecast.

Motivation

- Spotify is really making a music trend?
- Why many people uses Spotify's data?
- Are these reliable?

Analysis points

- Spotify is really making a music trend?
 - Compare with others' list
- Why many people uses Spotify's data?
 - Web scraping and using it
- Are these reliable?
 - ...

Data Sets by Web-scraping

- Spotify Weekly Top 100

(<https://spotifycharts.com/regional/us/daily/latest>)

- YouTube Music Chart Top 100

(<https://artists.youtube.com/charts/tracks>)

- Billboard the hot 100

(<https://www.billboard.com/charts/hot-100>)

iTunes
– for Genre

Spotify scraping

Scrapy

Spotify Charts

TOP 200 VIRAL 50 Filter by UNITED STATES DAILY 02/08/2018

TRACK	STREAMS
1 - God's Plan by Drake	3,348,966
2 - Stir Fry by Migos	1,306,069
3 - Mine by Bazzi	1,231,098

DOWNLOAD TO CSV

Elements Console Sources Network Performance Memory

```
<div class="responsive-select" data-type="country">...</div>
<div class="responsive-select" data-type="recurrence">...</div>
<div class="responsive-select" data-type="date">
  <div class="responsive-select-value">...</div> == $0
  <ul>
    <li data-value="2018-02-08" class="selected">02/08/2018</li>
    <li data-value="2018-02-07" class="selected">02/07/2018</li>
    <li data-value="2018-02-06" class="selected">02/06/2018</li>
    <li data-value="2018-02-05" class="selected">02/05/2018</li>
    <li data-value="2018-02-04" class="selected">02/04/2018</li>
    <li data-value="2018-02-03" class="selected">02/03/2018</li>
    <li data-value="2018-02-02" class="selected">02/02/2018</li>
    <li data-value="2018-02-01" class="selected">02/01/2018</li>
    <li data-value="2018-01-31" class="selected">01/31/2018</li>
    <li data-value="2018-01-30" class="selected">01/30/2018</li>
    <li data-value="2018-01-29" class="selected">01/29/2018</li>
    <li data-value="2018-01-28" class="selected">01/28/2018</li>
    <li data-value="2018-01-27" class="selected">01/27/2018</li>
    <li data-value="2018-01-26" class="selected">01/26/2018</li>
    <li data-value="2018-01-25" class="selected">01/25/2018</li>
    <li data-value="2018-01-24" class="selected">01/24/2018</li>
  </ul>
</div>
```

Spotify Charts

TOP 200 VIRAL 50 Filter by UNITED STATES DAILY 02/08/2018

TRACK	STREAMS
1 - God's Plan by Drake	3,348,966
2 - Stir Fry by Migos	1,306,069
3 - Mine by Bazzi	1,231,098

DOWNLOAD TO CSV

Elements Console Sources Network Performance Memory Application

```
</div>
<div class="loader hide"></div>
<table class="chart-table" style="display: table;" == $0
  <thead>...</thead>
  <tbody>
    <tr>
      <td class="chart-table-image">...</td>
      <td class="chart-table-position">1</td>
      <td class="chart-table-trend">...</td>
      <td class="chart-table-track">
        <strong>God's Plan</strong>
        <span>by Drake</span>
      </td>
      <td class="chart-table-streams">3,348,966</td>
    </tr>
    <tr>...</tr>
    <tr>...</tr>
    <tr>...</tr>
    <tr>...</tr>
  </tbody>
</table>
```

Selenium

The screenshot captures a web browser window displaying the YouTube Charts page for the US Top 100. The page features a navigation bar with tabs like CHARTS, MUSIC INSIGHTS, ARTIST RESOURCES, and BLOG. Below this, there's a section for VIRAL VIDEOS and ALL VIDEOS. The main content area lists the top 100 songs, with columns for TITLE, VIEW COUNT, CHANGE, and ON CHART. A date range dropdown menu is open on the left, showing various time periods. The right side of the browser window shows the Chrome DevTools interface, specifically the Console and Styles panels.

TITLE	VIEW COUNT	CHANGE	ON CHART
Bodak Yellow Cardi B	21.9M	0.8%	15 wks
Despacito (feat. Daddy Yankee) Luis Fonsi	14.7M	20%	39 wks
Look What You Made Me Do Taylor Swift	13.9M	20.7%	7 wks
I-800-273-8255 (feat. Alessia Cara & Khalid) Logic	13.1M	13.9%	8 wks
Rake It Up (feat. Nicki Minaj)	11.5M	11.7%	15 wks

Billboard scraping

Scrapy

Secure | <https://www.billboard.com/charts/hot-100/2018-02-10>

Apps BmarkTop A.Useful B.Kids C.US D.Sci D.Work E.ROK CMP Cloud Gmail GDrive Competitions canslove.tistory X-Speed Fin Forum Paper Uni Dc DEE W E

billboard

div.chart-row_main-display | 870 x 113

- Perfect
Ed Sheeran
- Finesse
Bruno Mars & Cardi B
- Havana
Camila Cabello Featuring Young Thug
- Rockstar
Post Malone Featuring 21 Savage

Elements Console Sources Network Performance Memory

```
<div class="chart-rail_module-video-wrapper">...</div>
<aside class="chart-sidebar-unit info-module highlights-module" data-trackaction=
"Highlights">...</aside>
<div class="chart-ad-ic">...</div>
<article class="chart-row chart-row--2 js-chart-row" data-hovertracklabel="Song
Hover-" data-songtitle">
  <div class="chart-row_primary">
    <div class="chart-row__history chart-row__history--steady">...</div>
    <div class="chart-row_main-display">...</div>
    <div class="chart-row__rank">
      <span class="chart-row__current-week">2</span>
      <span class="chart-row__last-week">Last Week: 2</span>
    </div>
    <div class="chart-row_image" style="background-image: url(https://charts-
static.billboard.com/img/2017/03/ed-sheeran-3vg.jpg)">
    </div>
  <div class="chart-row_container">...</div>
</article>
```

html body #main div div div article div.chart-row_primary div.chart-row_main-display

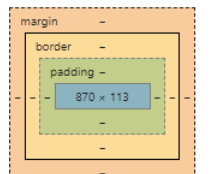
Styles Event Listeners DOM Breakpoints Properties Accessibility

Filter :hov .cls +

```
element.style {
}
.chart-row_main-display {
  overflow: hidden;
}
```

html, body, div, span, applet, charts.css?d399_480b714b7133:1
object, iframe, h1, h2, h3, h4, h5, h6, p, blockquote, pre, a,
abbr, acronym, address, big, cite, code, del, dfn, img, ins,

Console What's New



Why Spotify's dataset ?

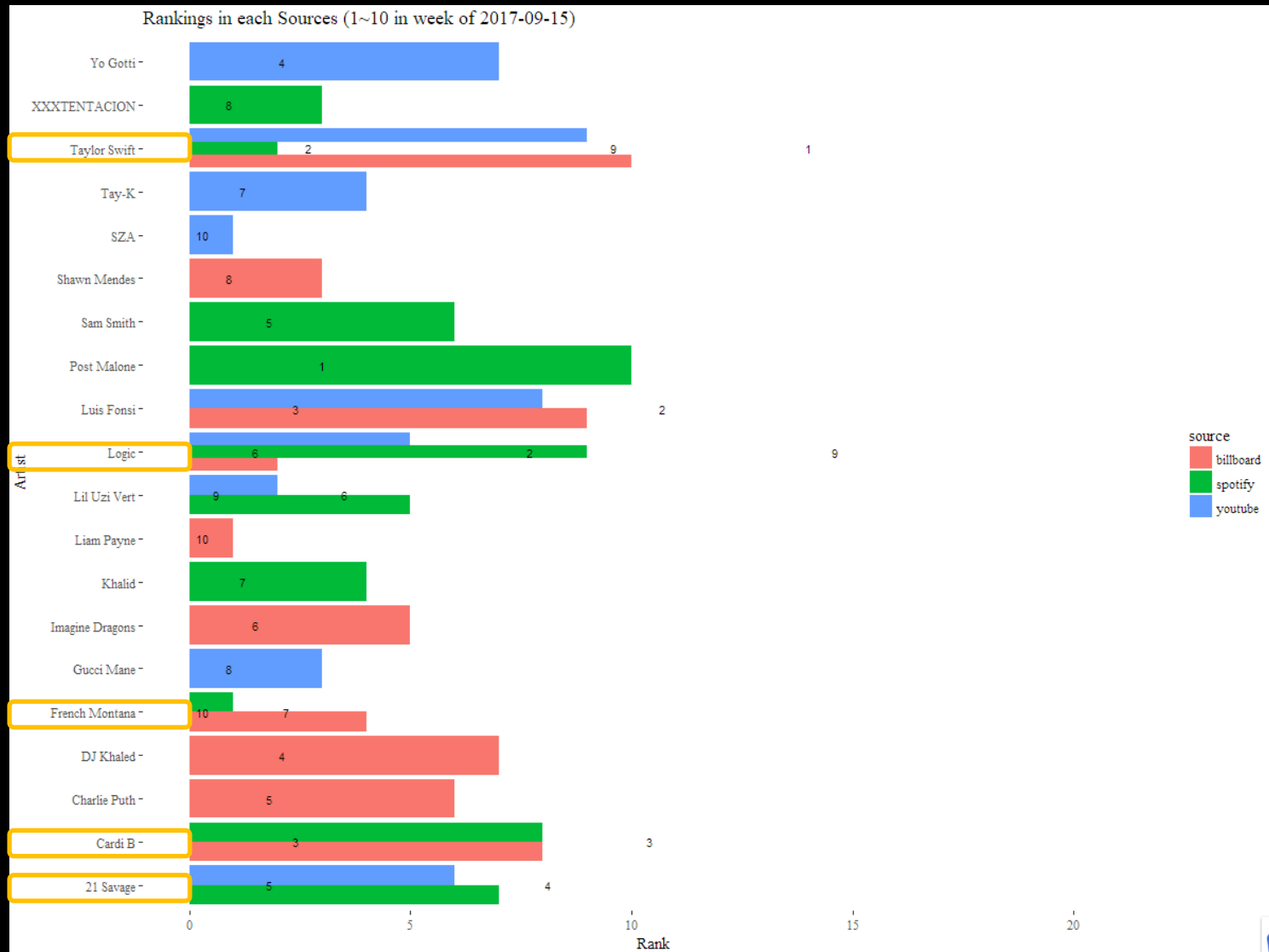
- Easy to web-scrape
- Easy to get (without web-scrape)
- Provide various data set
 - Regional, Daily, Weekly
- Long term monitored data
 - Every year
- Quick update

Rank Top10 Comparison

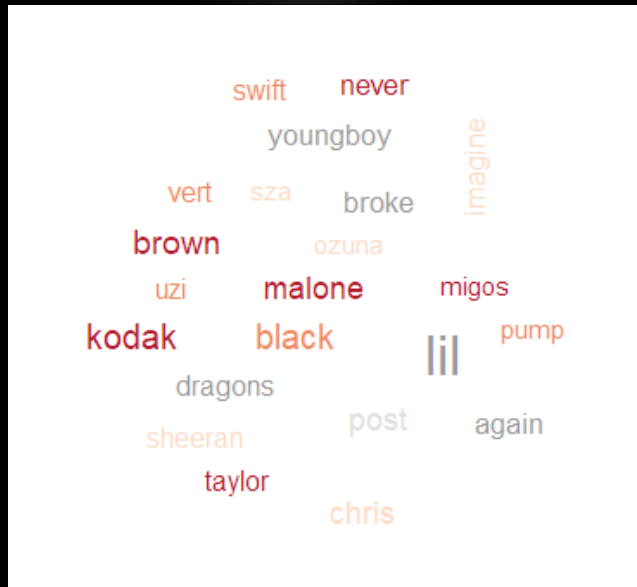
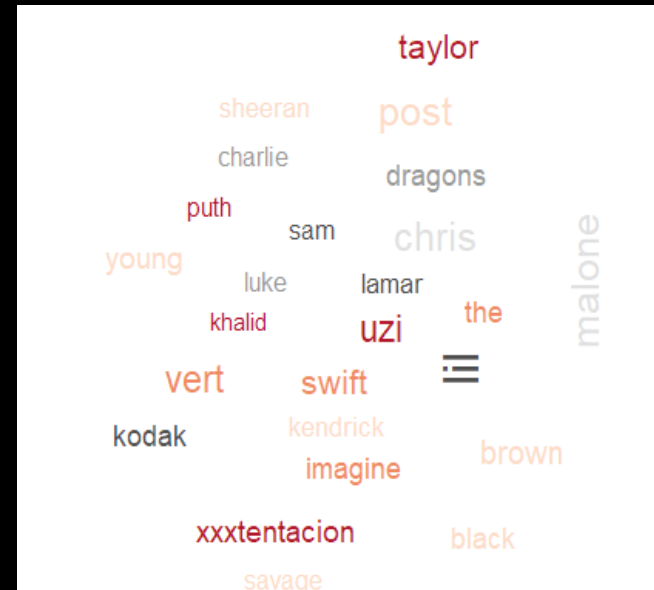
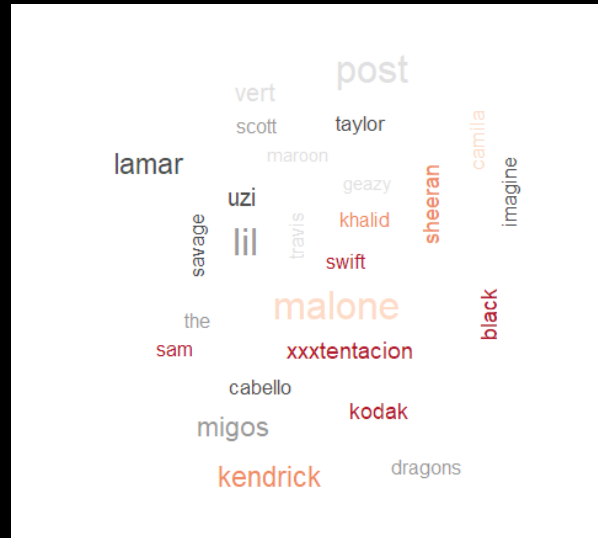
Rankings in each Sources (1~10 in week of 2017-09-15)



Rank Top10 Comparison



Significant factor selection



Significant factor selection

Artist who is registered in Top100 more than 30 times

Spotify	freq	Youtube	freq	Billboard	freq
Post Malone	110	lil ...	111	lil ...	97
lil ...	103	Kodak Black	57	Post Malone	70
Kendrick lamar	67	Brown	53	Chris	68
Migos	63	Chris	46	Uzi Vert	60
Uzi Vert	57	Post Malone	44	Taylor Swift	51
Kodak Black	47	Broke again	41	Brown	49
Ed Sheeran	45	Never youngboy	41	Young	44
Cabello Camila	42	Imagine Dragon	39	Kodak Black	42
Taylor Swift	41	Ed Sheeran	39	Imagine Dragons	40
Sam	39	Taylor Swift	38	Ed Sheeran	37
Imagine Dragon	38	Uzi Vert	37		

(2017-09-15 ~ 2018-01-26)

'Post Malone' : c("rockstar","congratulations","go flex","candy paint", "i fall apart", "white iverson")

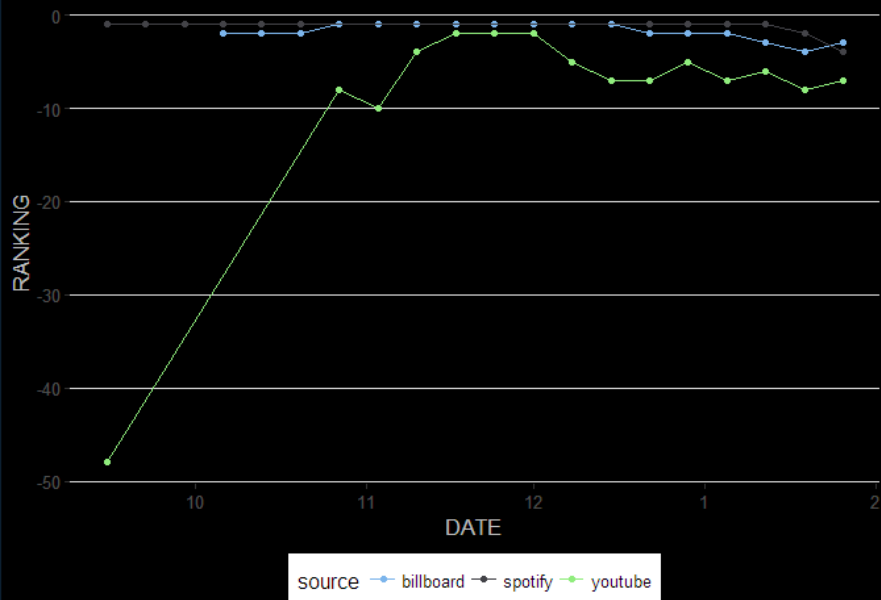
'Kodak Black' : c("roll in peace", "transportin", "tunnel vision", "codeine dreaming", "no flockin", "patty cake","halloween", "patty cake", "too much money", "roll in peace"))

'Taylor swift' : c("Look What You Made Me Do", "...ready for it?", "gorgeous", "call it what you want", "end game", "i did something bad", "delicate")

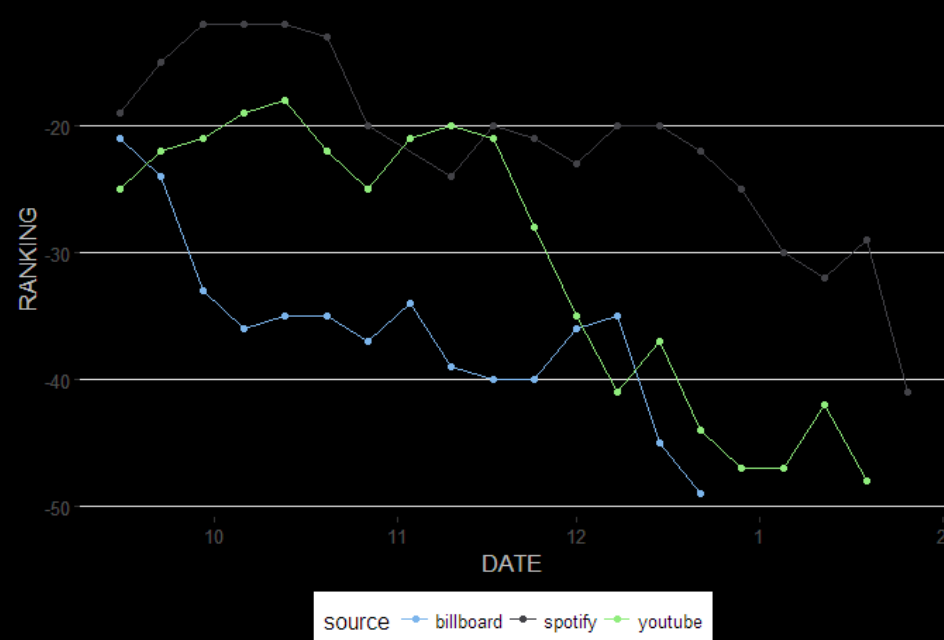
'Ed Sheeran' : c("shape of you", "perfect", "thinking out loud") # derivative should be merged again

'Imagine Dragons' : c("thunder", "believer")

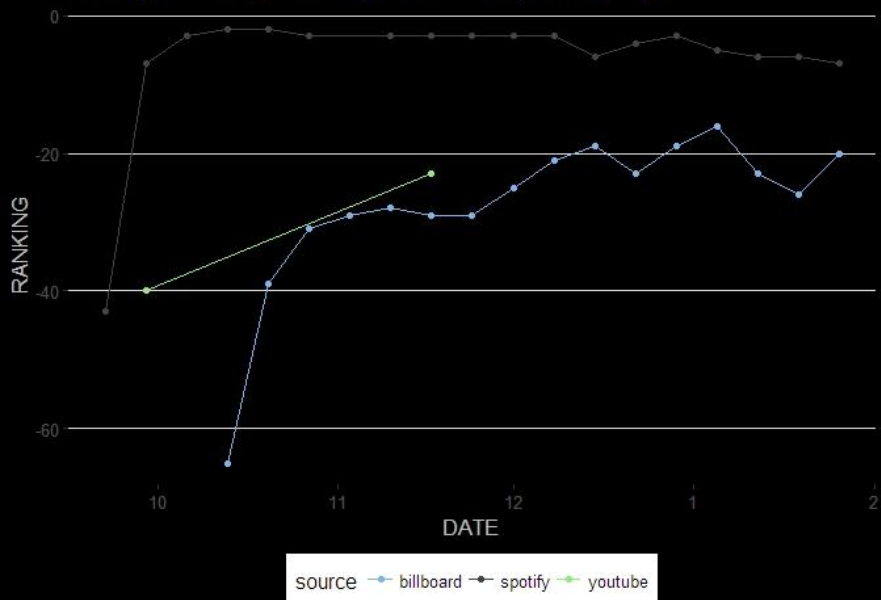
Weekly Ranking Trend - post malone(rockstar)



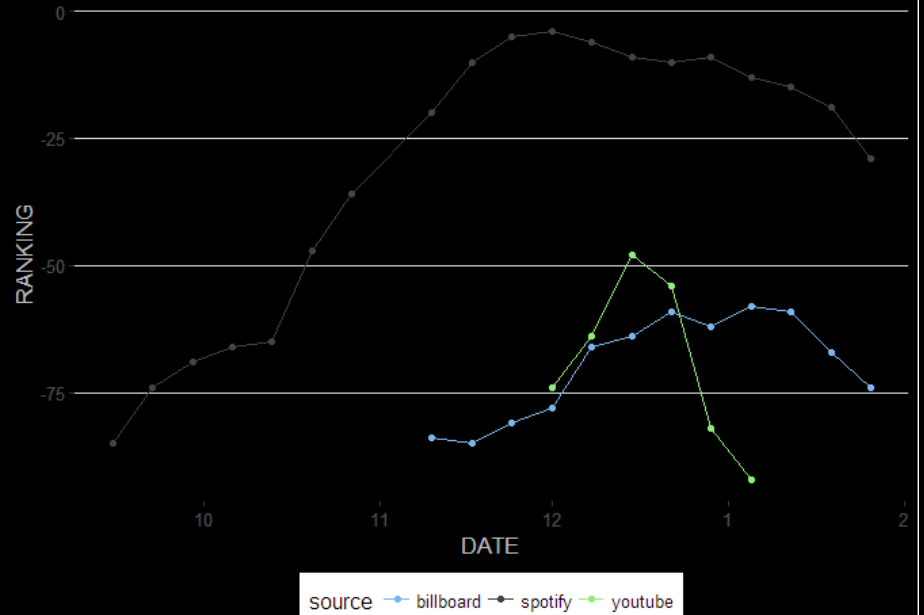
Weekly Ranking Trend - post malone(congratulations)



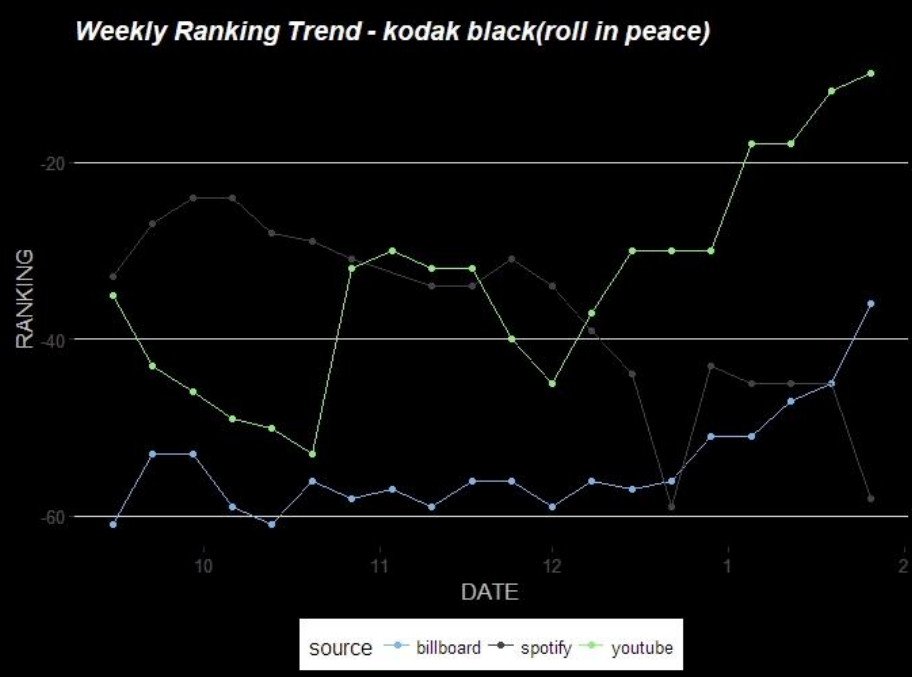
Weekly Ranking Trend - post malone(i fall apart)



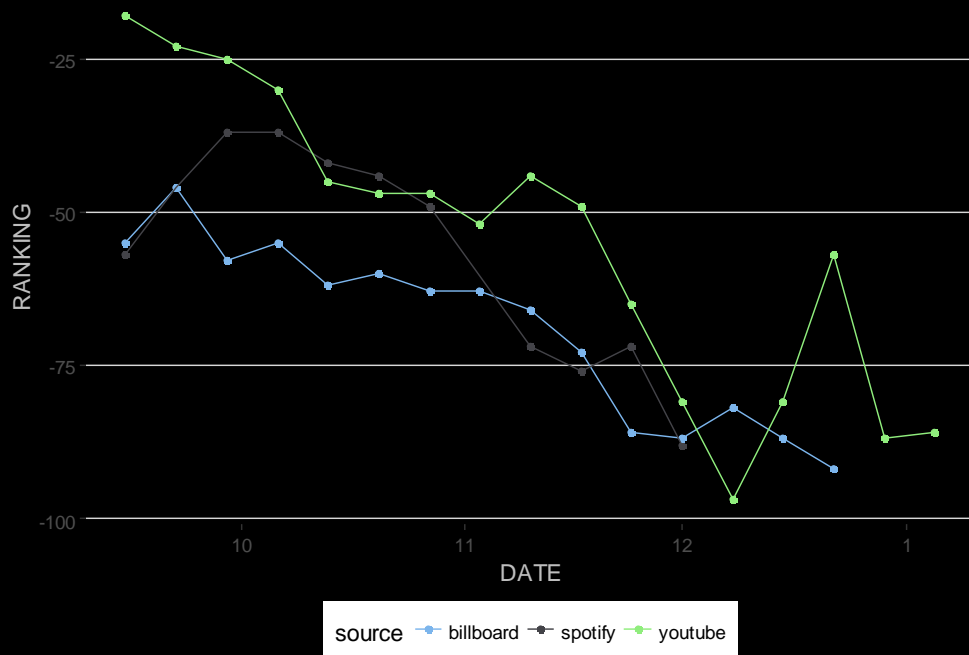
Weekly Ranking Trend - post malone(candy paint)



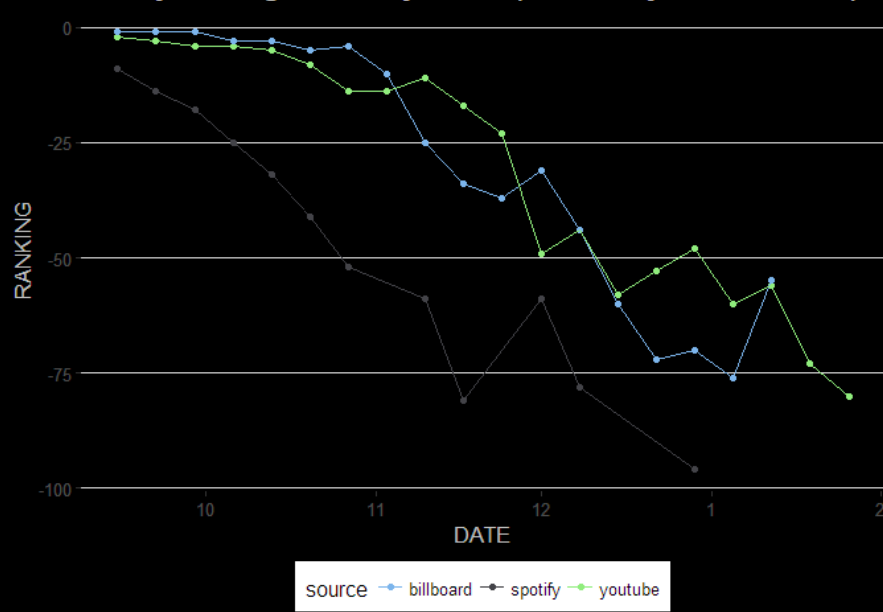
Weekly Ranking Trend - kodak black(roll in peace)



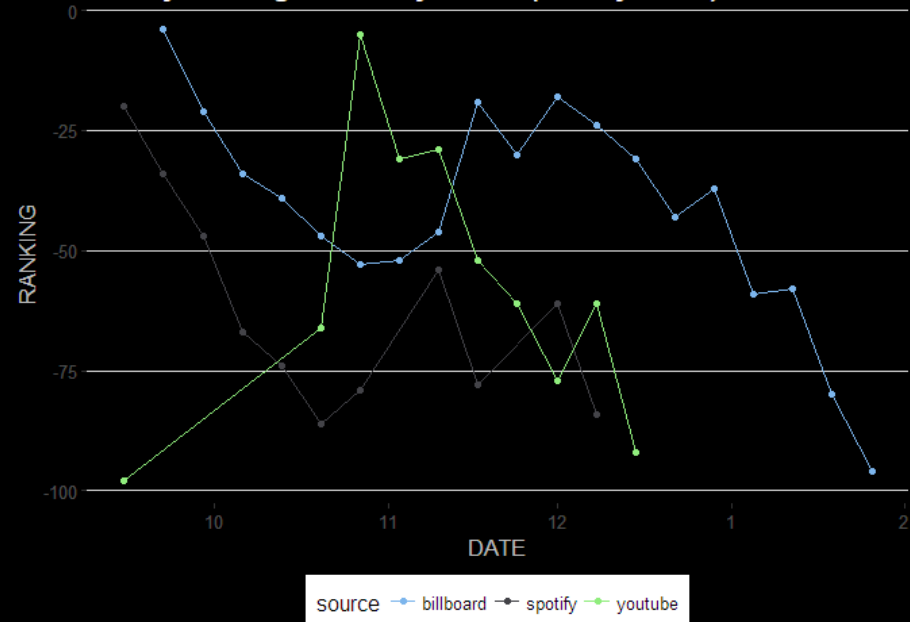
Weekly Ranking Trend - kodak black(transportin')



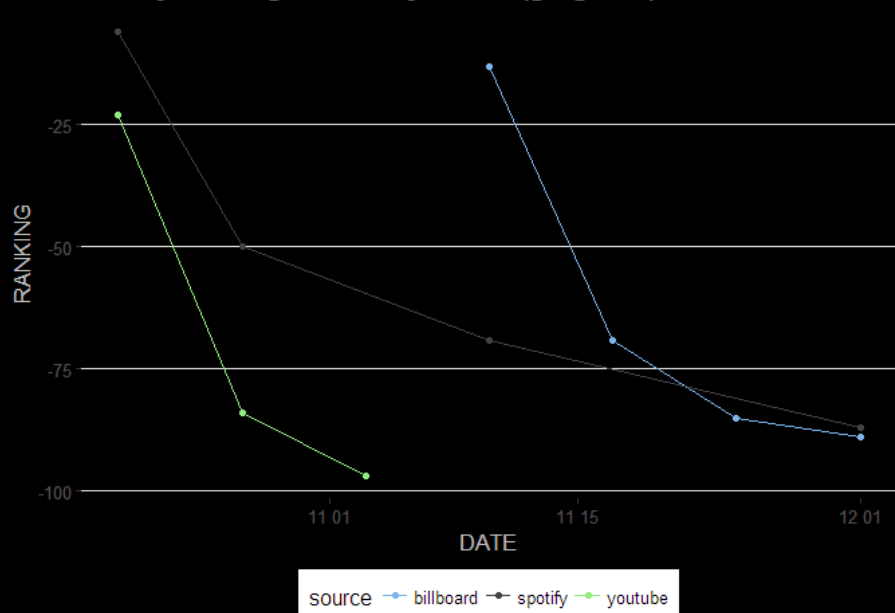
Weekly Ranking Trend - taylor swift(look what you made me do)



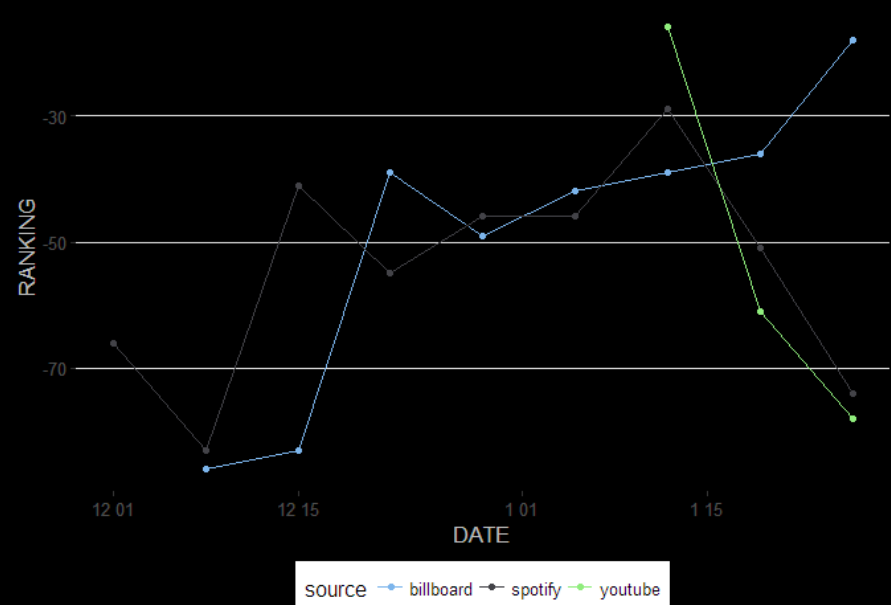
Weekly Ranking Trend - taylor swift(...ready for it?)



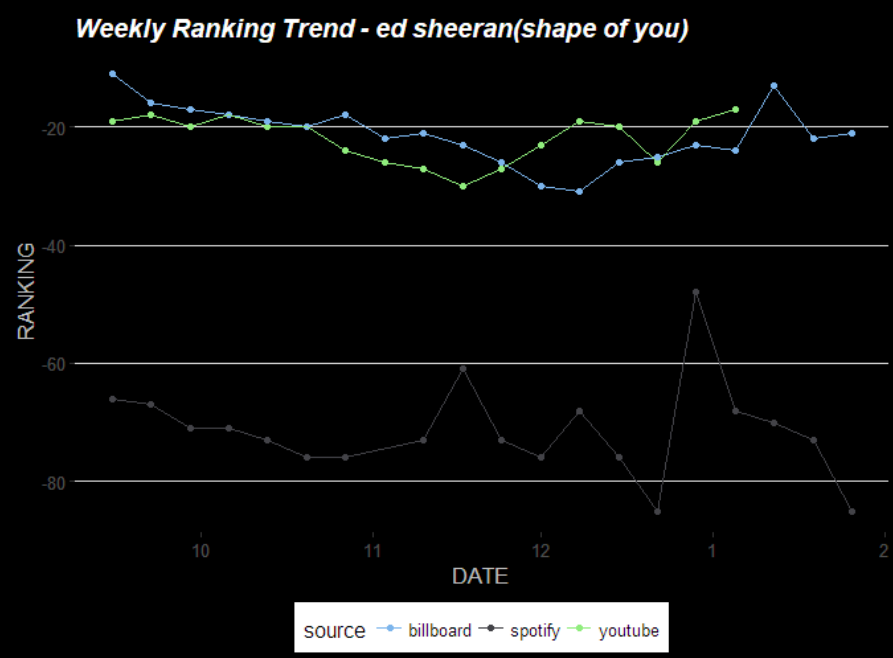
Weekly Ranking Trend - taylor swift(gorgeous)



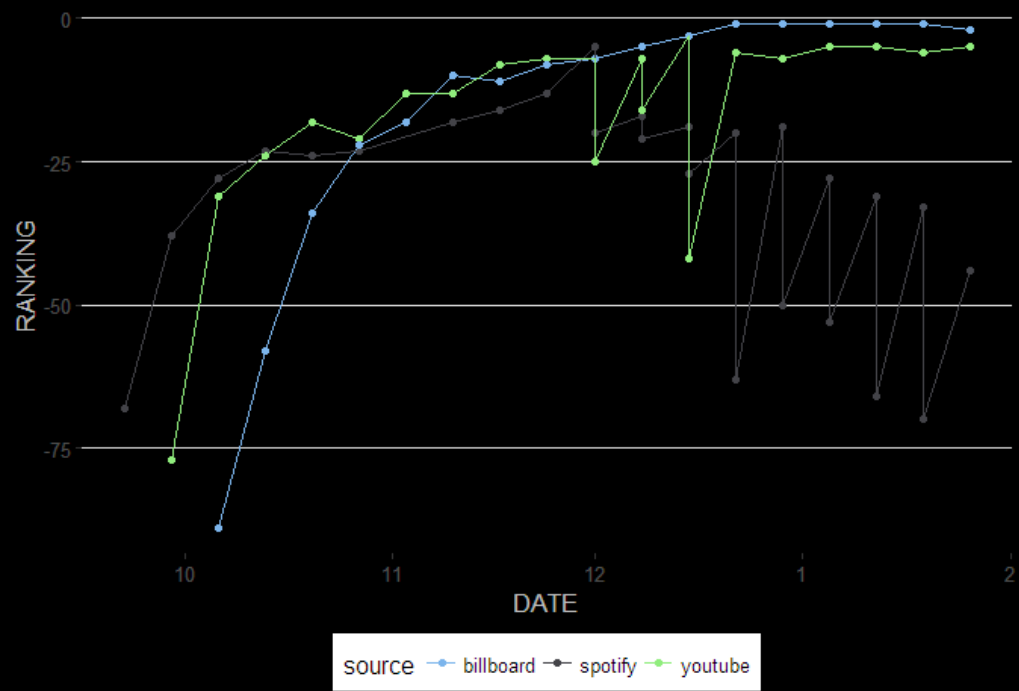
Weekly Ranking Trend - taylor swift(end game)



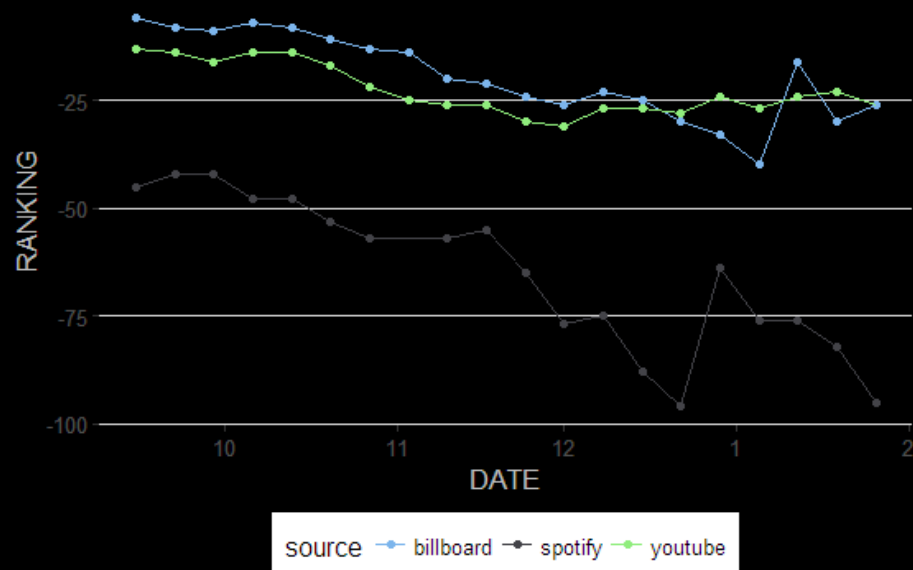
Weekly Ranking Trend - ed sheeran(shape of you)



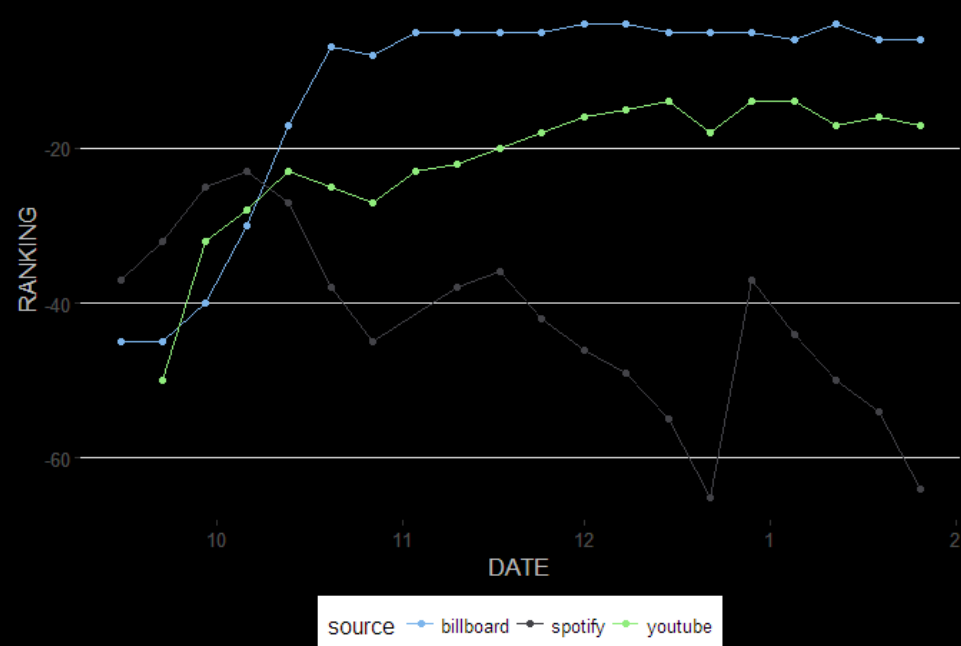
Weekly Ranking Trend - ed sheeran(perfect)



Weekly Ranking Trend - imagine dragons(believer)



Weekly Ranking Trend - imagine dragons(thunder)



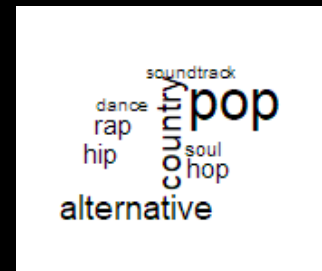
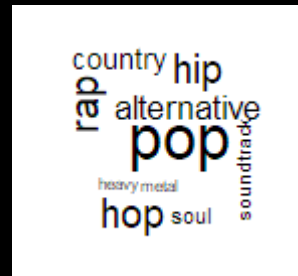
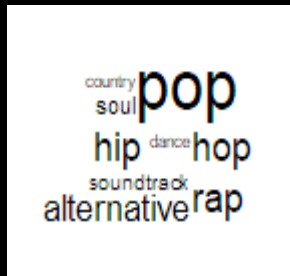
Observation

- Spotify's ranking line is on the left-side than YouTube music and Billboard's
- Spotify provide first viral hit and followed by YouTube music and Billboard's

→ But,

long-term data comparison is needed

Genre positioning in Top100



Genre frequency in Week Top100 (2017-09-15 ~ 2018-01-26)

Spotify		Youtube		Billboard	
pop	49.6%	pop	44.0%	pop	40.5%
hip hop/ rap	21.2%	hip hop/rap	21.4%	country	18.6%
alternative	18.7%	alternative	15.7%	alternative	18.3%
R&B/Soul	5.7%	country	11.1%	hip hop/rap	11.9%
Soundtrack	4.2%	R&B/Soul	4.7%	R&B/Soul	4.5%
Country	0.2%	Soundtrack	2.9%	Soundtrack	3.6%
Dance	0.2%	Heavy metal	0.2%	Dance	2.7%

How to Rank



Total number of
downloaded stream



Simple,
but
Fast, and effective !



Total number of
viewed music on
YouTube by video,
artist, track, and
viral score



Radio airplay
audience
impressions as
measured by
Nielsen Music,
sales data as
compiled by
Nielsen Music and
streaming activity
data provided by
online music
sources

Future Work

- Data Survey
 - Regional subscriber distribution
 - Weather, Airplay data
- Analysis
 - Free vs. charged subscriber
 - By age, gender, occupation
 - By weather, date
- Visualization

**Thank you
for
attention !**