Reading Trends using Reddit postings

- Natural Language Processing

SJ Choi

Questions

1. How much distinguishable reddit posts are?
     r/HeadphoneAdvice vs r/headphones

2. Can we find trends in Headphone by reddit posts?
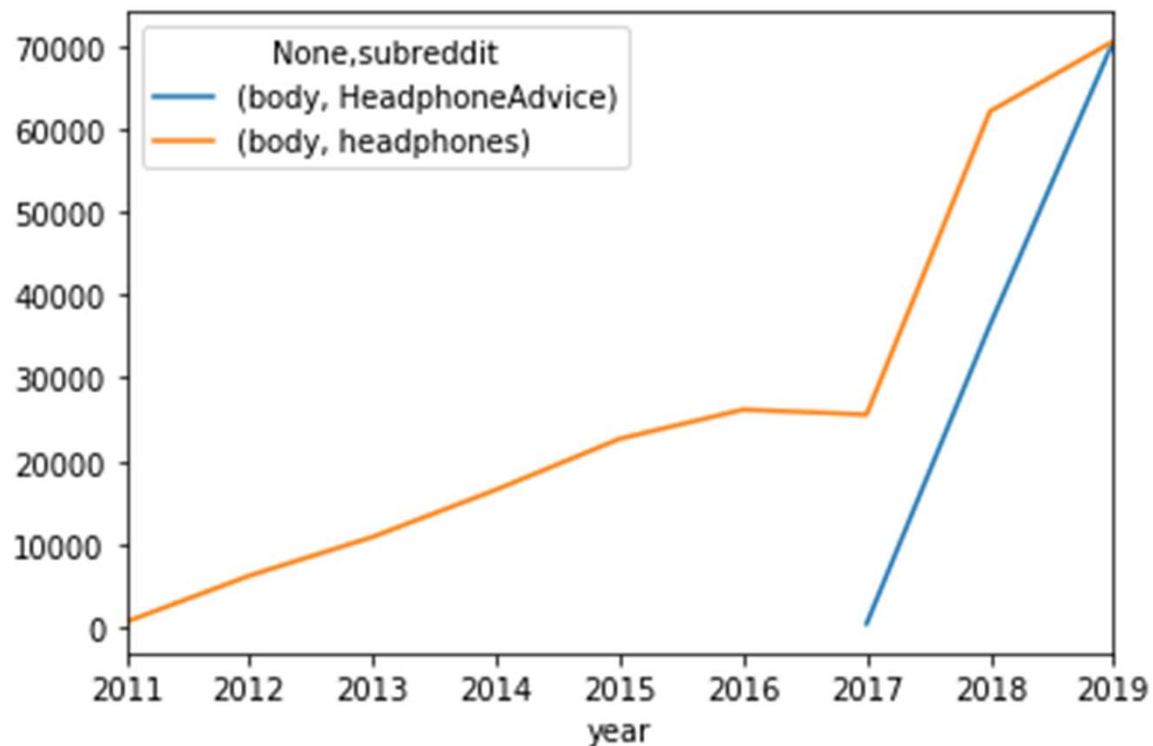
r/HeadphoneAdvice vs. r/headphones
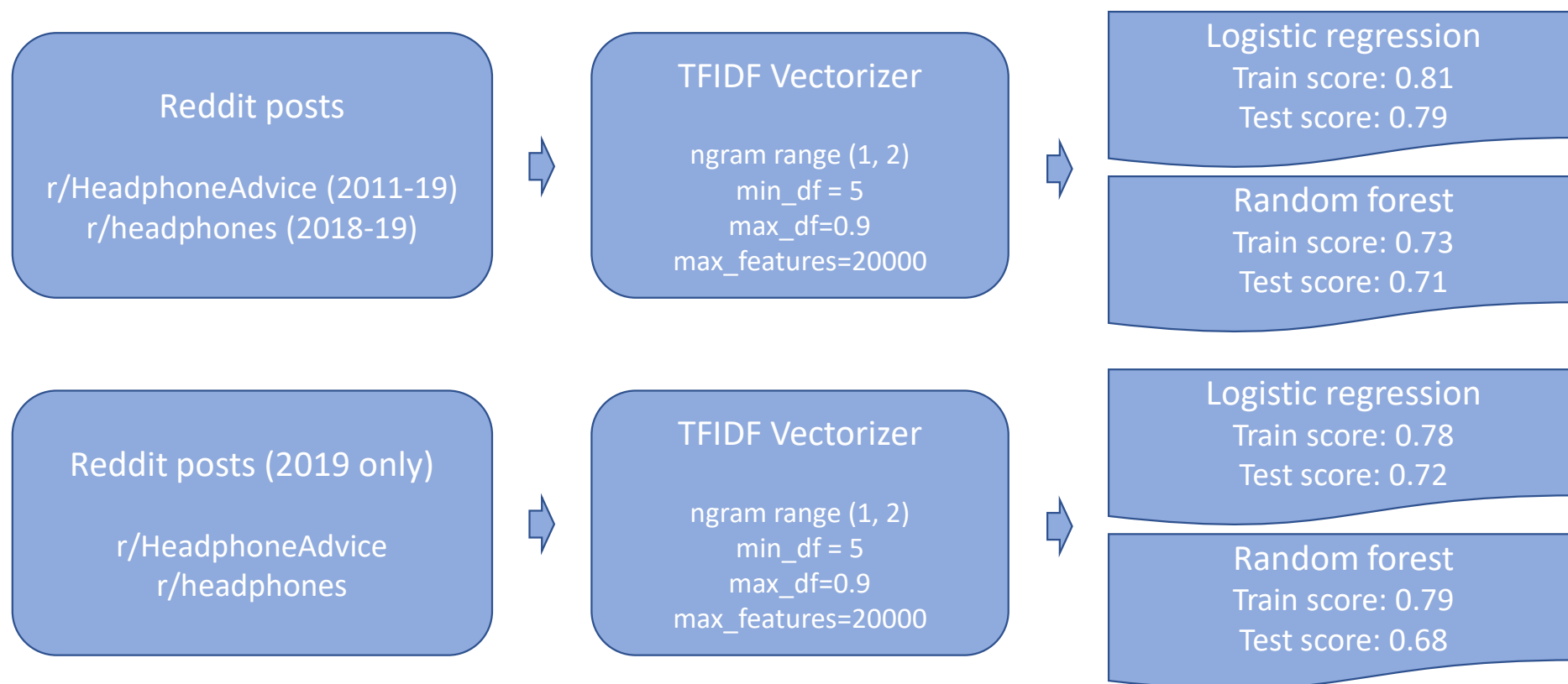
(feat. Regression & Random forest)

As r/headphones was launched in 2017, number of postings of both subreddits were increased together.

Modeling – Logistic regression & Random forest

- As the test scores stay below 0.8, we can say r/HeadphoneAdvice and r/headphones share similar contents.

Reddit posts

r/HeadphoneAdvice (2011-19)
r/headphones (2018-19)

→

TFIDF Vectorizer

ngram range (1, 2)
min_df = 5
max_df=0.9
max_features=20000

→

Logistic regression
Train score: 0.81
Test score: 0.79

Random forest
Train score: 0.73
Test score: 0.71

Reddit posts (2019 only)

r/HeadphoneAdvice
r/headphones

→

TFIDF Vectorizer

ngram range (1, 2)
min_df = 5
max_df=0.9
max_features=20000

→

Logistic regression
Train score: 0.78
Test score: 0.72

Random forest
Train score: 0.79
Test score: 0.68

Reading Trends of Headphone by reddit posts

(feat. LDA* modeling & visualization)

*LDA: Latent Dirichlet Allocation

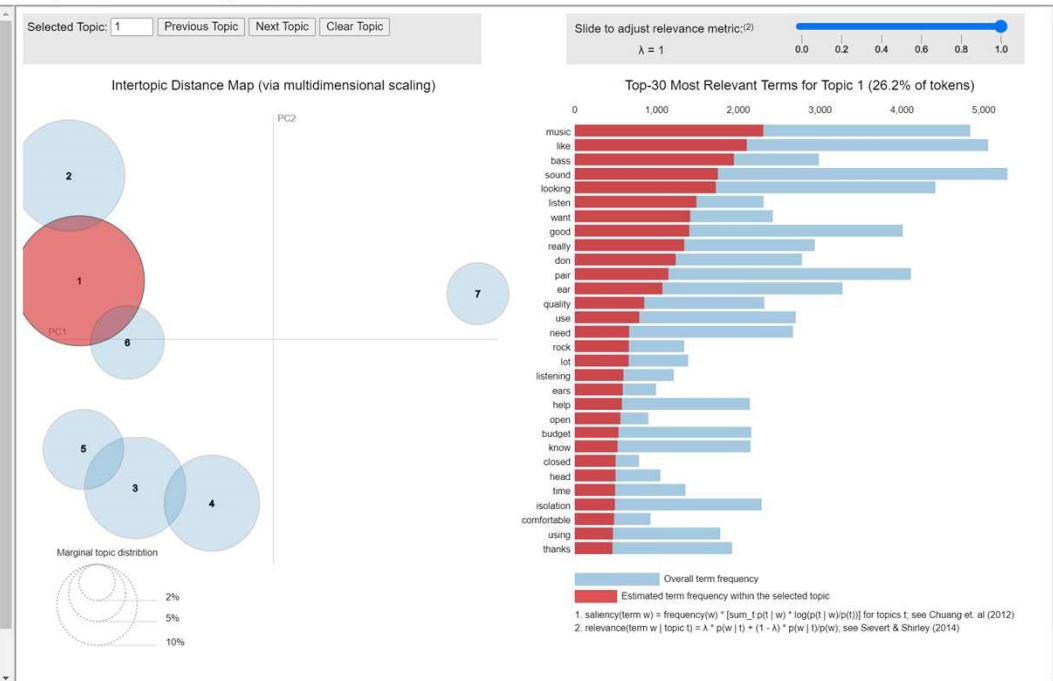# LDA Modeling – 2013 data
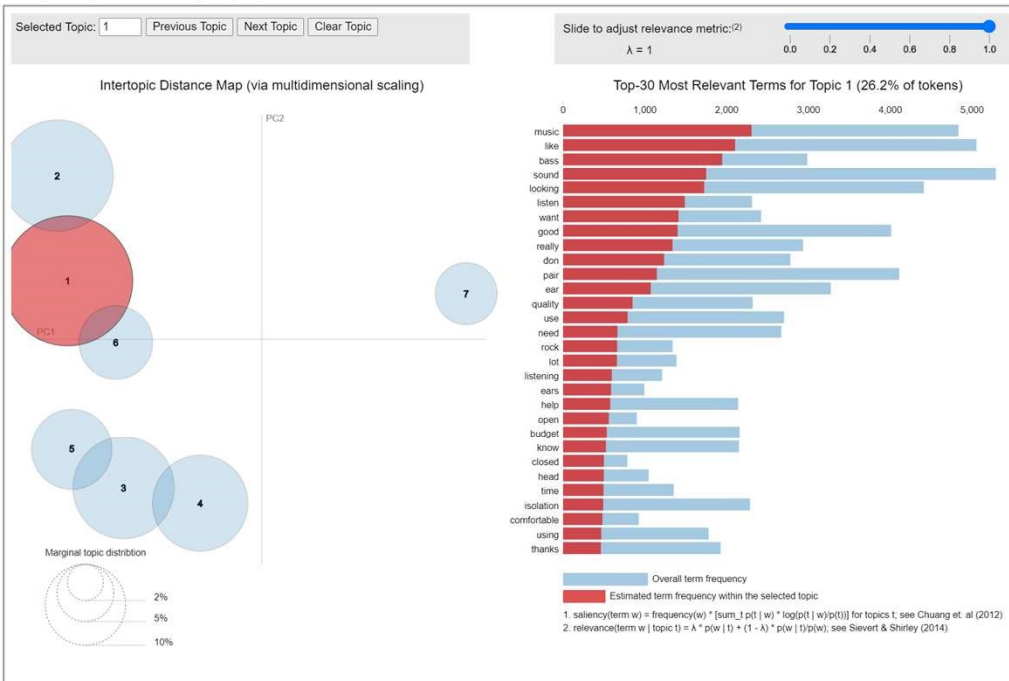
2013 - 7 topics from 10.9 comments

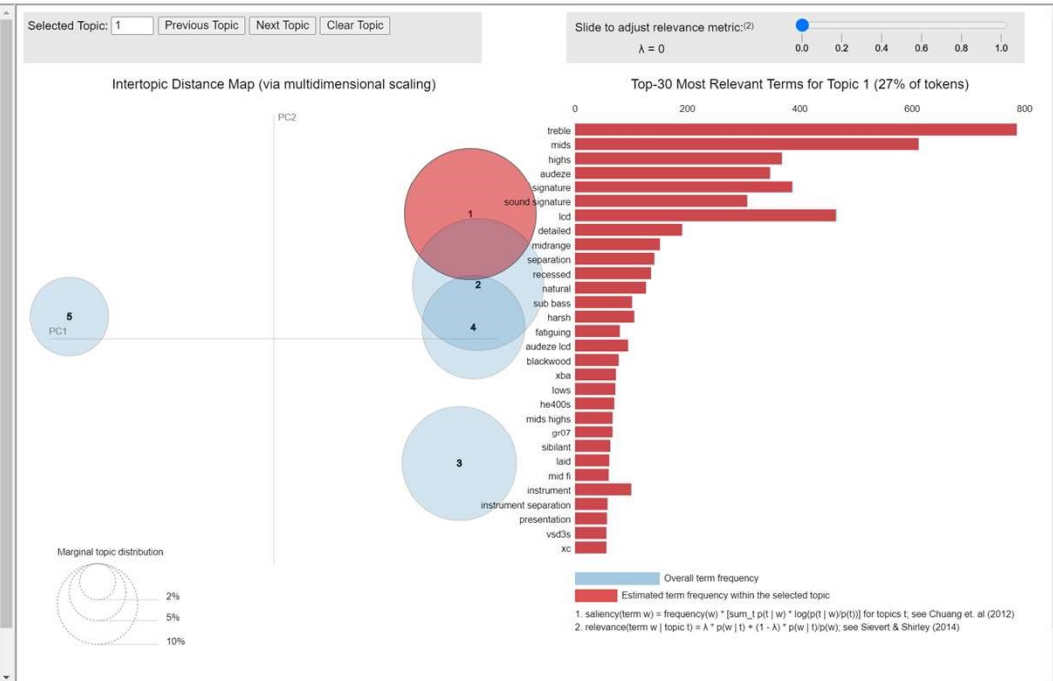| Topic	(%)	Overall terms	Topic specific terms
Listening to music by genre	26.2	Music, like, bass, looking, sound, good, pair, ears, quality, rock, help, budget, closed, head	bass don, fi buying, heavy bass, lust, mad lust, sure look, gaming guide, gaming read, rock, quality, electronic, hip hop
Balance for watching youtube	19.5	Preferred, music, isolation, budget, looking, past, type, source, like, balance, preferred music, pair, preferred type, tonal, tonal balance, youtube, watch	Youtube watch, requirements isolation, preferred tonal, preferred type, watch, type sized, ear preferred, balanced,
Amp sound	15.9	Sound, like, pair, good, bass, got, new, price, better, think, really, know, bit, head, little, great, cans, beats, akg, Sennheiser	Impressions, hd800, k701, tubes, audeze, bottlehead, sibilance, tube amps,
beyerdynamic DT 770 PRO	14.1	Amp, audio, dac, sound, ath, fiio, m50, use, pro, dt, need, card, beyerdynamic, ath m50, audio technical	Dt 770, 770 pro, beyerdynamic, dt990, e17, pro250
Earbuds	10	Ear, pair, new, cable, earbuds, help, know, right, need, looking, like, best, quality, left, use	Wire, earbud, left ear, right ear, pair earbuds, stopped working, foam tips, ear canal, way fix, good earbuds
Gaming headset	8.4	Headset, gaming, sound, Sennheiser, mic, looking, good, wireless, hd, pc,	Gaming headset, wireless, surround sound, steelseries siberia, plantoronics, headset mic g930
amp	5.9	Amp, amazon, dp, sr, ref, utf8, sr, qid, ath	Utf8 amp, ref, qid, amp sr, amp qid, amp keywords, gp

LDA Modeling – 2015 data
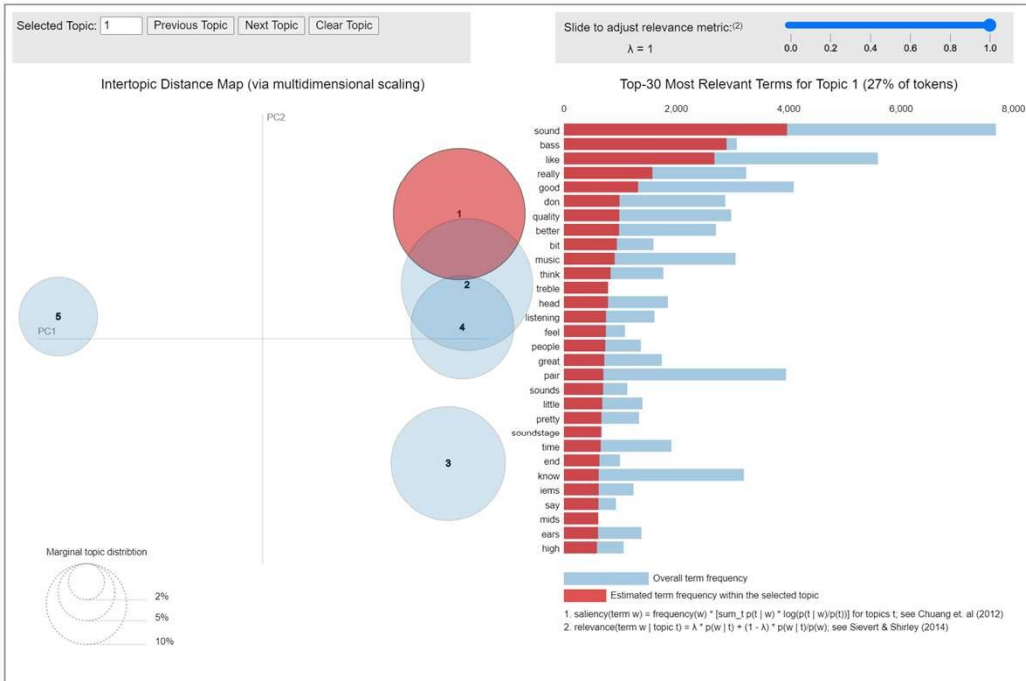
2015 - 5 topics from 22.7k comments

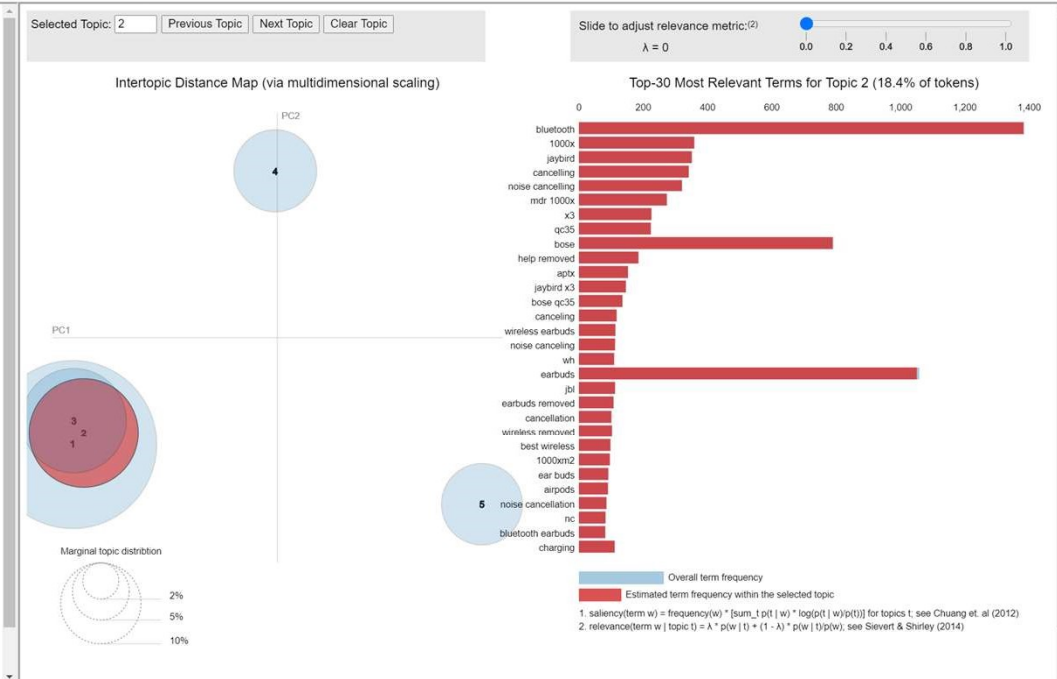| Topic	(%)	Overall terms	Topic specific terms
Sound quality for listening to music	27	Bass, sound, like, really, good, treble, soundstage, mids, bit, quality, don, think, know, high, little, better, music	Treble, mids, highs, audeze, sound signature, lcd, detailed, midrange, separation, recessed, natural, sub bass, instrument
Wireless earphone issues	26.8	Ear, pair, cable, like, help, know, left, right, got, don, use, sound, problem, noise, bought, pads	Bose, repair, noise cancelling, stopped working, ear buds, Kraken, black Friday, ear cup, cushion
Amp, dac	20.3	Amp, dac, sound, audio, Amazon, power, gaming, headset, use, need, using, Sennheiser, good, USB, Fiio, looking, like, volume, mic	Ref, sound card, utf8 amp, surround, poll, qid, amp sr, amp qid, xonar, asus
Recommenda tions	16.4	Ath, music, HD, Sennheiser, audio, good, m50x, Audio technical, Schiit, removed, listen, Youtube	Ath m50x, uber, Magni modi, crossfade, HD650, AKG k7xx
Gaming headset	9.5	Guide, looking, buying, fi, amp, head fi, alternatives, gaming, check, beats, read, posting, make sure, google, comments	Buying guide, alternatives, spreadsheet, docs, advice thread, looking gaming, looking budget, faq,

LDA Modeling – 2017 data
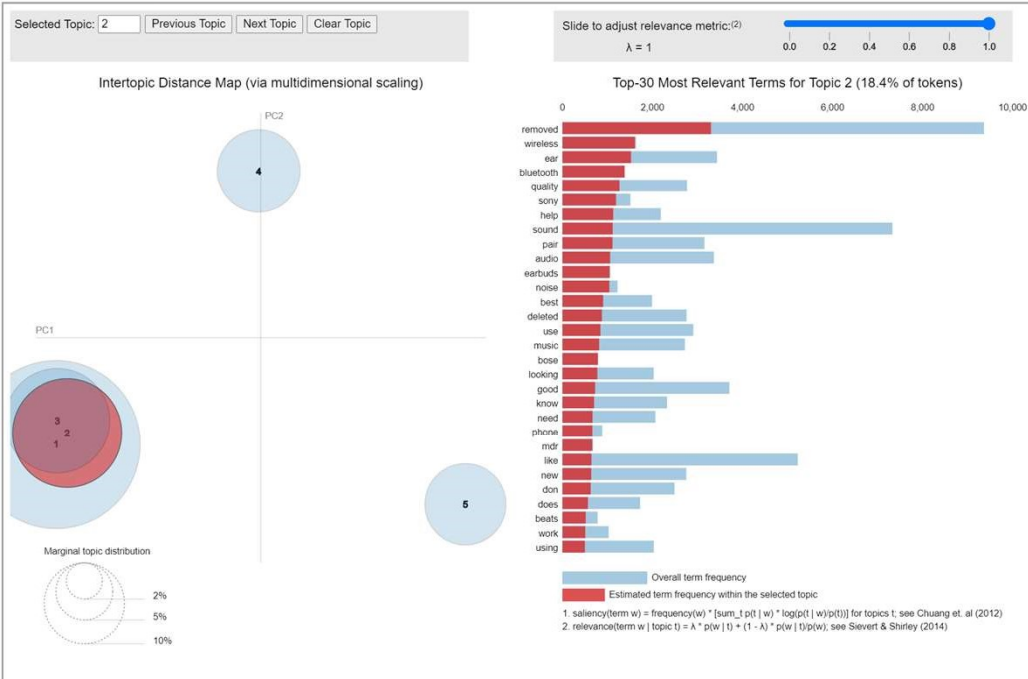
2017 - 5 topics from 26.0k comments

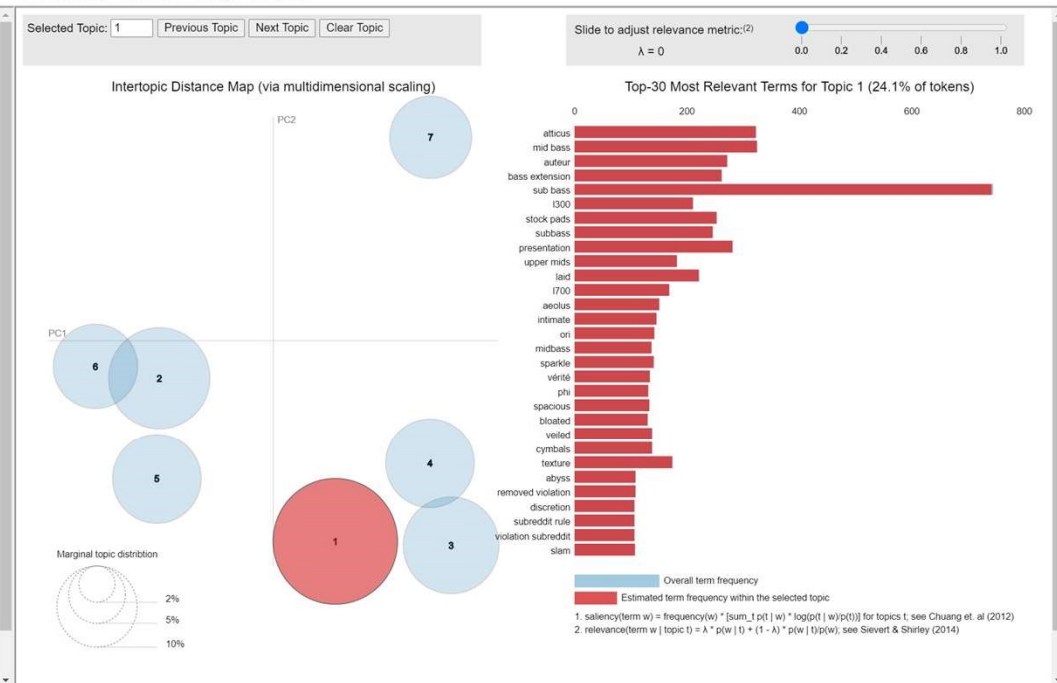| Topic	(%)	Overall terms	Topic specific terms
Sound quality for listening to music	43.8	Sound, like, don, really, good, better, music, amp, bit, pads, listening, bass, feel	Bass, treble, sound signature, vocals, mids, soundstage, warm, highs, detailed, clarity
Noise cancelling & earbud	18.4	Removed, wireless, ear, Bluetooth, quality, Sony, help, sound, pair, audio, earbuds, noise, best, deleted, use, looking	Bluetooth, Bose, earbuds, 1000x, jaybird, noise cancelling, x3, qc35, help removed
Gaming headset	17	removed, Sennheiser, deleted, cable, HD, new, gaming, good, pair, headset, know, looking, Amazon, got	Fc, hz gain, pk fc, gain db, 00 filter, hyperx cloud, gaming headset, filter pk, apo
Amp, dac, Tech support	10.6	Dac amp, removed, audio, tech support, questions, need, thread, impedance, amplifier, output	Tech support, support questions, sound card, week discussion, damage equipment
Purchase advice	10.1	Purchase advice, thread, head fi, question, guide, google, assistants, spreadsheet, comments, wiki	Purchase advice, guides, assistants, spreadsheet wiki, purchase help, answer question, requesting

LDA Modeling – 2018 data
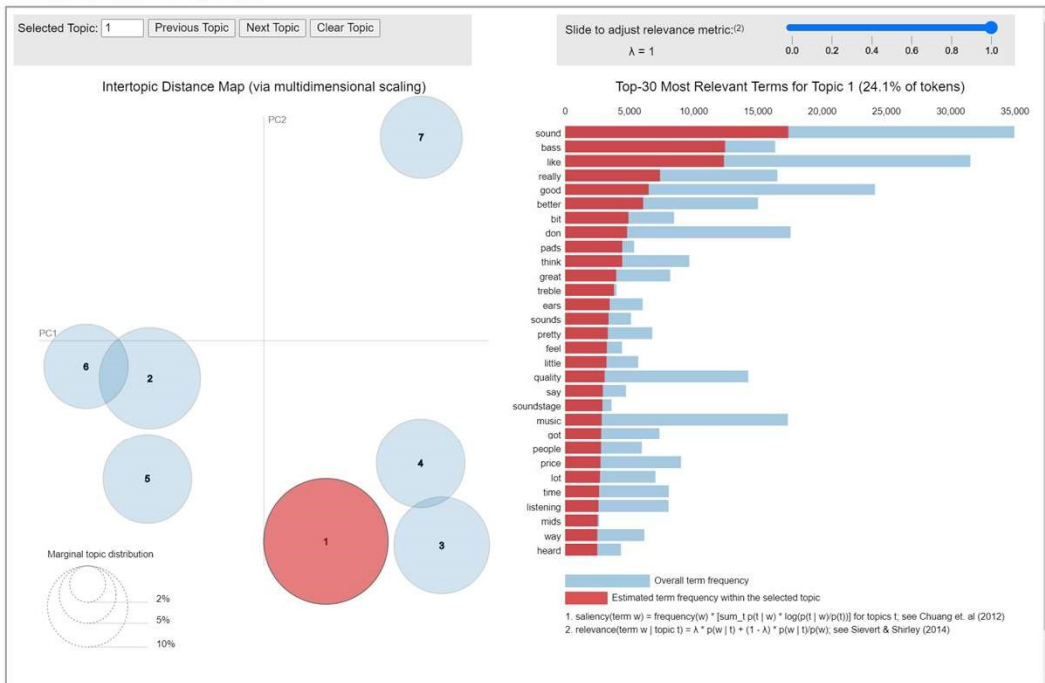
2018 - 7 topics from 98.5 comments

Topic	(%)	Overall terms	Topic specific terms
High-end headset	30	Sound, like, really, bass, good, don, think, better, bit, people, got, way, time, right, audio, treble	Utopia, elex, 2c, lcd 2c, l300, dropbox, dl, sound different, pdf dl, tyll, l700, innerfidelity, auteur, mastered, flange, resolving
Noise cancelling & earbud	19.2	Wireless, ear, Bluetooth, sony, noise, looking, good, sound, quality, pair, earbuds, like, bose, best, use, sound quality, noise cancelling	Anc, sony wh, noise cancelling, qc35, battery life, wireless earbuds, airpods, Jabra, jaybird, 1000xm2, xm3, active noise
Listening to music by genre	17.7	Music, bass, like, sound, listen, looking, budget, good, listening, open, want, amp, pa*	Tonal balance, genre, preferred music, past gear, hip hop, gear experience, source amp, edm, like improve, rap, experience like
Gaming headset	17.3	Amp, gaming, audio, mic, headset, good, Sennheiser, sound, looking, use, need, hd, amazon, pc, ath, pair	Hyperx cloud, gaming headset, utf8 amp, amp sr, amp qid, amp keywords, modmic, xbox, arctis
Amp, dac, Cable	15.8	Amp, dac, cable, audion, pads, use, need, removed, power, schiit, deleted, usb, know, massdrop, impedance	Deleted, d30, rca, q1, hz gain, fc, gain db, pk, d30, tech support, replacement pads, balanced cable

*PA: A public address (PA) system is for amplifying the human voice. In its simplest form, it has a microphone, mixer, and loudspeakers.
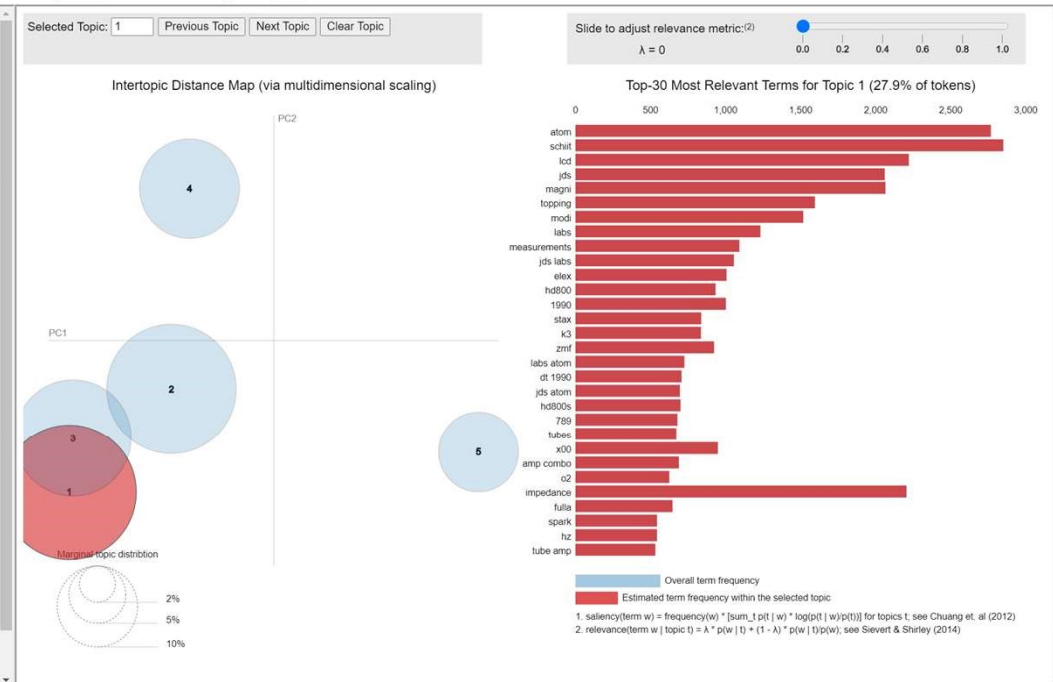
LDA Modeling – 2019 data
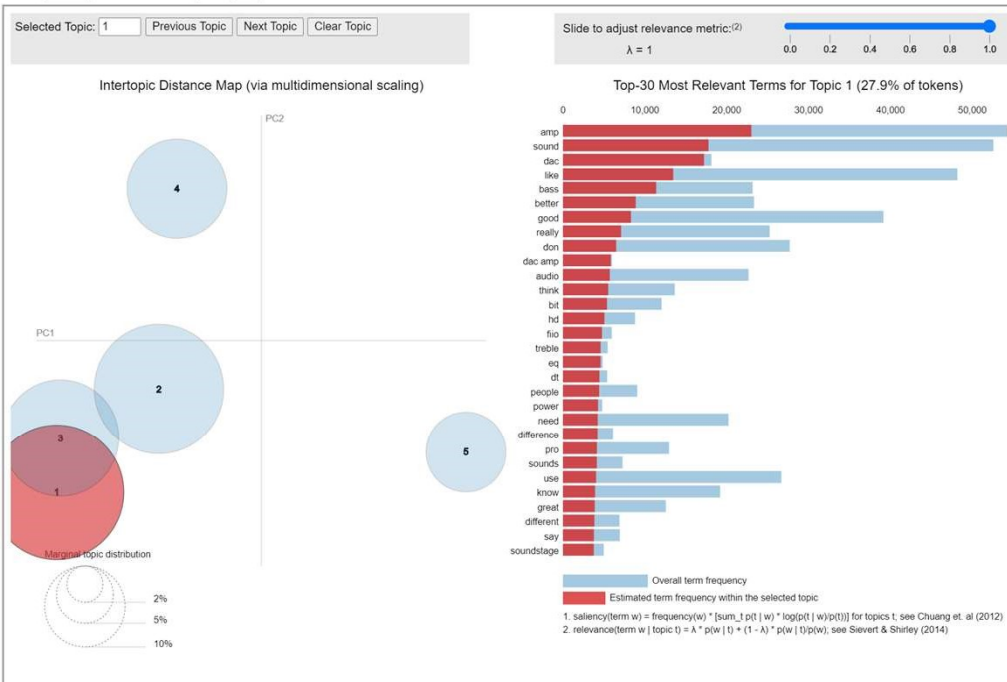
2019 - 5 topics from 141.0k comments

Topic	(%)	Overall terms	Topic specific terms
Sound Quality with DAC AMP	27.9	Dac, amp, sound, bass, EQ, treble, power Fiio, DT, like, 6xx, difference, HD, soundstage, bit	Atom, Schiit, LCD, JDS labs, Magni, topping, Modi, impedance
Listening to music by genre	25.9	Music, looking, budget, listen, don, like, pair, bass, good, PA*, really, want, experience, used	Tonal balance, genre, past gear, hip hop, preferred tonal, gear experience, preferred music, rap, like improve, tin audio, tin t2
Ear buds	20.9	Earbuds, like, wireless, don, cable, battery, tips, fit, airpods, sound, good, pair, know, really, got, right, post, think, use, quality, tried	Wireless earbuds, 65t, Jabra elite, galaxy buds, impressions review, submit, powerbeats, Canal, clicking, charging case
ANC** Headset	15.4	Amp, x200b, Sony, Bluetooth, noise cancelling, wireless, Bose, sound, quality, audio, Amazon, Sennheiser, good, usb, ANC, looking	REF, dp, amp qid, amp sr, amp psc, gp, pxc Sony wh, psc, Bose qc 35,
Gaming headset	9.9	Gaming, headset, mic, looking, removed, thanks, good, need, cable, help, advice, buy, pro	hyperx cloud, rule gt, following rule, sites, google, google view, view quipa, quipa assistants, purchase assistant, quick recommendation, mk2

*PA: A public address (PA) system is for amplifying the human voice. In its simplest form, it has a microphone, mixer, and loudspeakers.
**ANC: Active Noise Cancelling

Trend of categories

- Similar categories keep discussed.

Listening to music	Amplifire, DAC, other peripherals	Earbuds & ANC**	Gaming headset	Purchase advice	Others

2013 (10.9k*)	(%)	**2015** (22.7k*)	(%)	**2017** (26.0k*)	(%)	**2018** (98.5k*)	(%)	**2019** (141.0k*)	(%)
Listening to music by genre	26.2	Sound quality for listening to music	27	Sound quality for listening to music	43.8	High-end headset	30	Sound Quality with DAC AMP	27.9
Balance for watching youtube	19.5	Wireless earphone issues	26.8	Noise cancelling & earbud	18.4	Noise cancelling & earbud	19.2	Listening to music by genre	25.9
Amp sound	15.9	Headset for PC usage	20.3	Gaming headset	17	Listening to music by genre	17.7	Ear buds	20.9
beyerdynamic DT 770 PRO	14.1	Recommendations	16.4	Amp, dac, Tech support	10.6	Gaming headset	17.3	ANC** Headset	15.4
Earbuds	10	Gaming headset	9.5	Purchase advice	10.1	Amp, dac, Cable	15.8	Gaming headset	9.9
Gaming headset	8.4								
amp	5.9								

*Number of comments
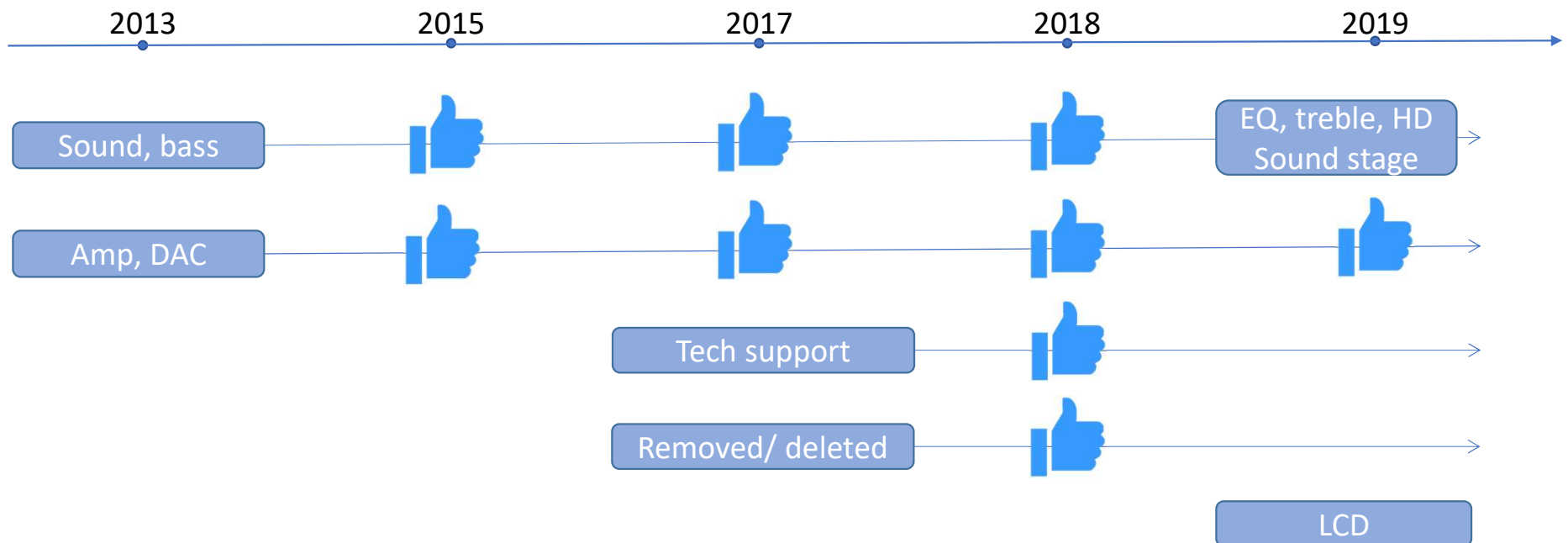**ANC: Active Noise Cancelling

Mainly discussed factors – Listening to music

- Sound fit for preferred music genres is one of the important factors considered when people select a headphone.
- The quality of expression of bass & treble is a popular theme for listening to music.
- Tonal balance also becomes an important factor as well as bass or treble itself.
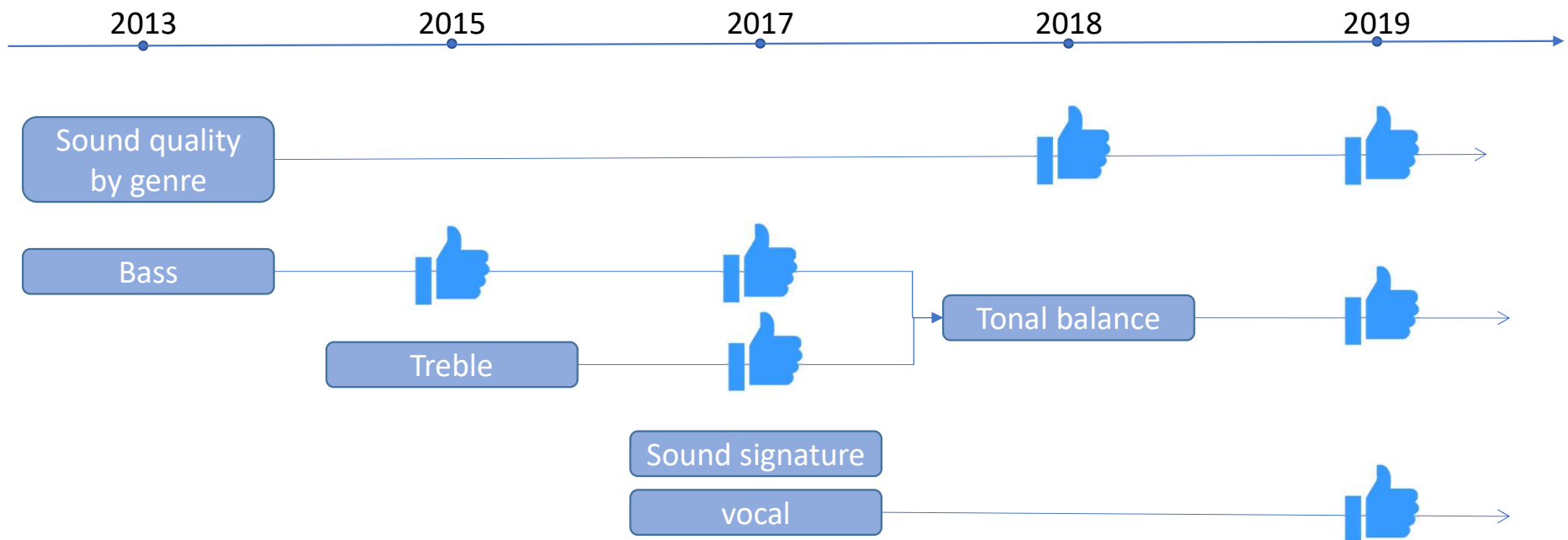
Mainly discussed factors – Amplifier, DAC

- This is an unfamiliar category for most consumers and some people feel they need technical support.
- Discussed sound quality was evolved from bass to EQ, treble, HD or sound stage.

Mainly discussed factors – Earbuds & ANC**

- Earbuds were discussed first then, noise cancelling / wireless / battery life were added.
- Active Noise Cancelling headphones became another main stream along with earbuds.
- Good looking has been always important and battery is a concerned point.
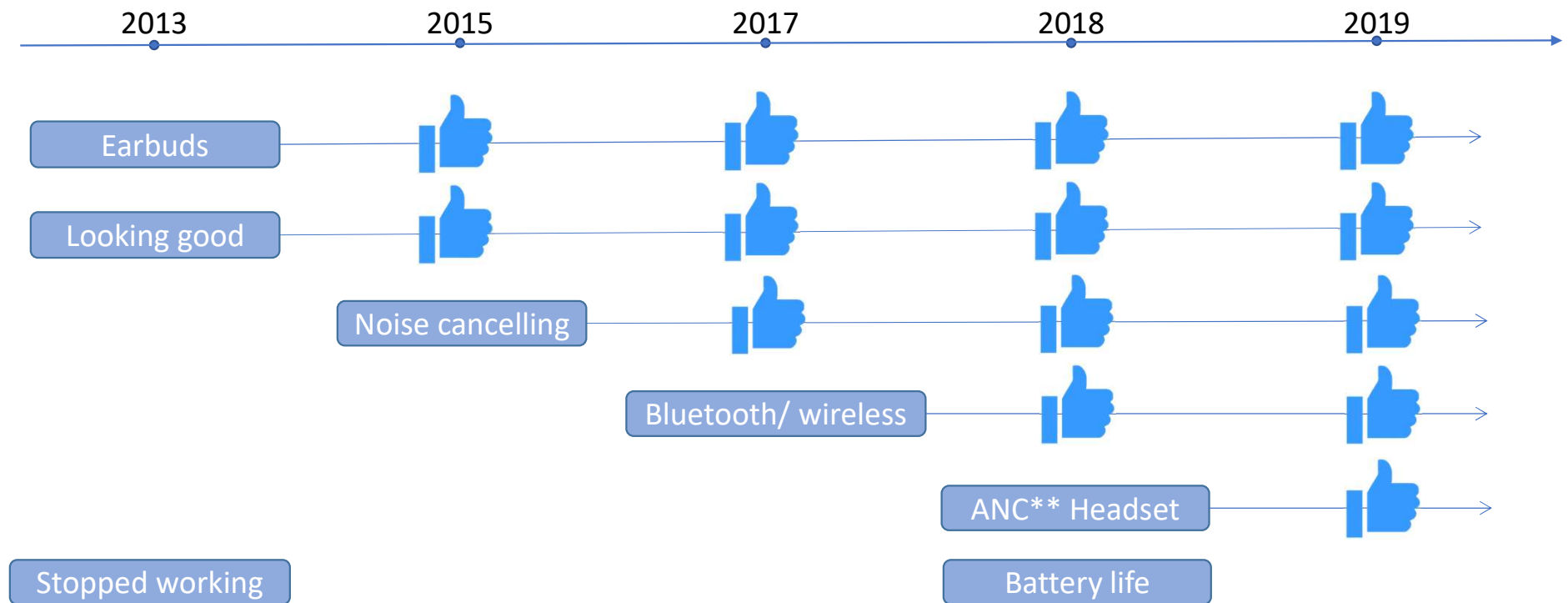
**ANC: Active Noise Cancelling

Mainly discussed factors – Gaming headset

- Main important factors are constant: good looking, microphone, PC connection, optimized sound effect.
- Brand-wise, gaming headset has a different category from earbuds or active noise cancelling headphones.
- Based on the terms in this category, gamers are more techy than usual consumers.
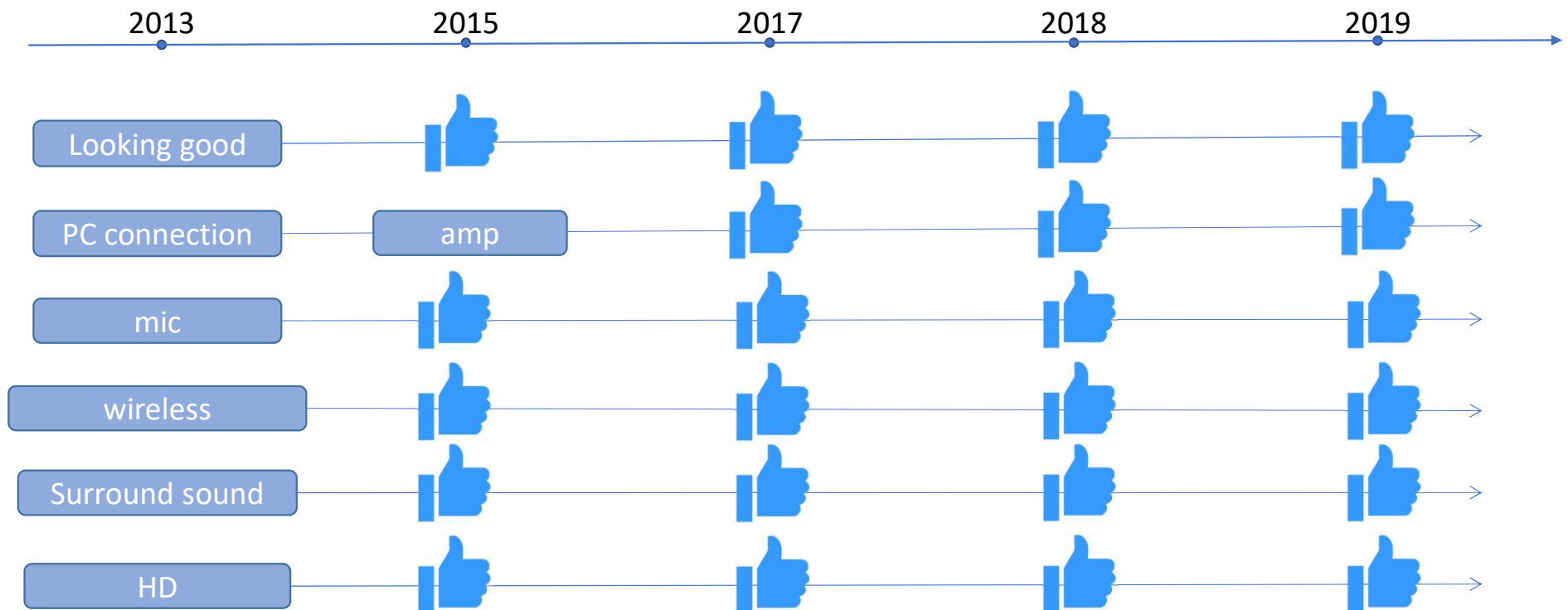
Executive summary

1. Since 2017, the number of postings about headphones has skyrocketed and earbuds & noise cancelling headsets have played a key role.

2. Sound quality in listening to music moved its focus from bass & treble to tonic balance. Also preferred music genre is an important factor for selecting a headphone.

3. Usual consumers are not familiar with amplifier, DAC or other peripheral devices for better sound experience, which can be an opportunity for marketing(ex. Word of mouth), manufacturing(ex. Intuitive UI) or sales(ex. Bundling.)

4. Earbuds and Active noise cancelling headset became two separate and conspicuous main streams: the TPO* of them may be different but the consumers can be the same.

5. Gaming headset is undeniably a unique mainstream and the techy users constantly pursue better looking and better performance.

*TPO: Time, Place, Opportunity

Thank you!