

Exploring music evolution over the last century



Data Wrangling 2024-2025, Jan 25th



Radu Nicolae

BSc Computer Science
@atLarge Research
@Network Institute,



Leonid Ivanov

@VU Amsterdam,
@atLarge Research



**Alara
Karadeniz**
@VU Amsterdam



Ayush Khadka
@VU Amsterdam



Human Nature
Michael Jackson



Il bacio di Klimt
Emanuele Aloia



The Real Slim Shady
Song • Eminem



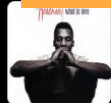
Erik Dalı / Sendemi Oldun Ankaralı / Huriyem
Omer Faruk Bostan

Student Talent Award '25,
Faculty of Science

WHAT IS MUSIC?

a pattern of sounds made by musical instruments, voices, or computers, or a combination of these, **intended to give pleasure to people listening to it.**

Source: Cambridge



What Is Love - 7" Mix
▶ Music video • Haddaway

Music is everywhere



decisions

=> health :D



 = productivity impact

 = health impact

 = psychological impact

Sources

- [1] How Music Helps People Heal: The Therapeutic Power of Music, Victor et al.
- [2] The sound of stress recovery, Springer 2023, Adiasto et al.
- [3] The digital transformation of societies Music, United Nations
- [4] Neural correlates of music perception and Cognition, Nature Neuroscience

TODAY'S AGENDA

1. ~~What is music~~
2. Research Methodology (MRQ + RQs)
3. Data Sources (Kaggle + Spotify API)
4. Data Wrangling Methods – from bytes to plots,
5. Findings – answering (M)RQ(s)
6. Threats to validity (limitations)
7. Let's conclude

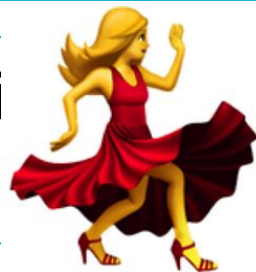
RESEARCH METHODOLOGY

Background

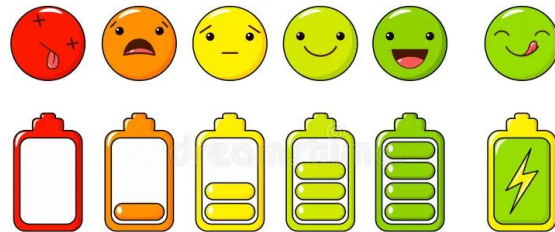
Acousticness indicator of whether the song is acoustic or contains classical instruments



Danceability a collection of different factors indicating how suitable a given track is for dancing



Energy perceptual measure of a song's intensity



Duration...



Main research question

MRQ: How did music evolve over the last century?

RQ1: How did the acousticness-energy ratio in songs change over the last century?

RQ2: How did the energy-danceability ratio in songs change over the last century?

RQ3: How did the duration of songs evolve over the last century?

DATA SOURCES

Methods

Data collection

Data Extraction and Processing

Data Creation

Data Analysis

Data Collection

Kaggle and Spotify API



Data Extraction and Processing

mode	year	acousticness	danceability	duration_ms	energy	instrum
1	1921	0.8868960000000005	0.4185973333333336	260537.16666666663	0.23181513333333334	0.34487
1	1922	0.9385915492957748	0.4820422535211267	165469.74647887325	0.23781535211267596	0.4341
1	1923	0.9572467913513516	0.5773405405405401	177942.36216216214	0.2624064864864865	0.37173
1	1924	0.940199860169493	0.5498940677966102	191046.70762711862	0.3443466101694912	0.5817
1	1925	0.9626070503597138	0.5738633093525181	184986.92446043165	0.2785935251798561	0.4182
1	1926	0.660817216981134	0.5998802612481859	156881.65747460088	0.2114670907111756	0.3330
1	1927	0.9361794552845558	0.6482682926829262	184993.59837398372	0.2643213008130081	0.3913
1	1928	0.9386165035685952	0.5342878667724027	214827.90642347344	0.20794779540047573	0.4948
1	1929	0.6014265861344558	0.6476698529411761	168999.41281512607	0.2418007352941172	0.21520
1	1930	0.936714937370057	0.5181758835758836	195150.28534303536	0.3335239189189189	0.35220
1	1931	0.8330399585921336	0.5952217391304357	171553.42546583852	0.2344967732919257	0.22141

Data Creation & Data Analysis

	year	energy	acousticness	duration_sec	danceability	energy_acousticness_ratio	energy_danceability_ratio
0	1921	0.232	0.887	260.537	0.419	3.823276	1.806034
1	1922	0.238	0.939	165.470	0.482	3.945378	2.025210
2	1923	0.262	0.957	177.942	0.577	3.652672	2.202290
3	1924	0.344	0.940	191.047	0.550	2.732558	1.598837
4	1925	0.279	0.963	184.987	0.574	3.451613	2.057348
...
95	2016	0.593	0.284	221.397	0.600	0.478921	1.011804
96	2017	0.590	0.286	211.116	0.612	0.484746	1.037288
97	2018	0.602	0.268	206.001	0.664	0.445183	1.102990
98	2019	0.593	0.278	201.025	0.645	0.468803	1.087690
99	2020	0.631	0.220	193.728	0.693	0.348653	1.098257

▸ Data wrangling

▸ Data acquisition

▸ Data cleaning

▸ Data merging

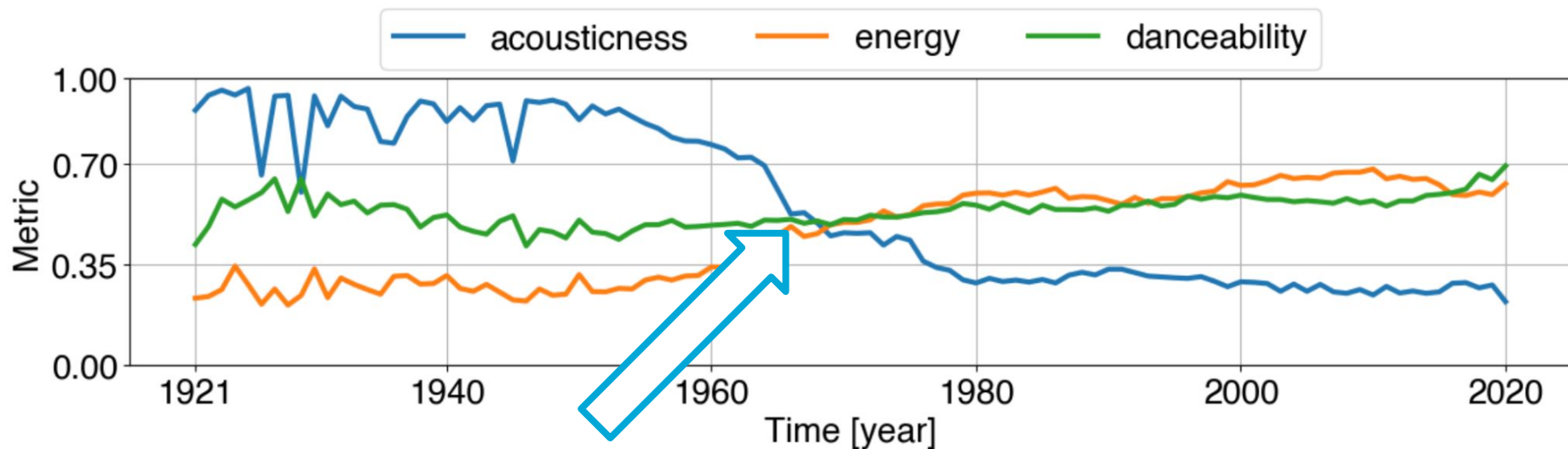
▸ Data aggregation

▸ Data visualization

▸ Time series data

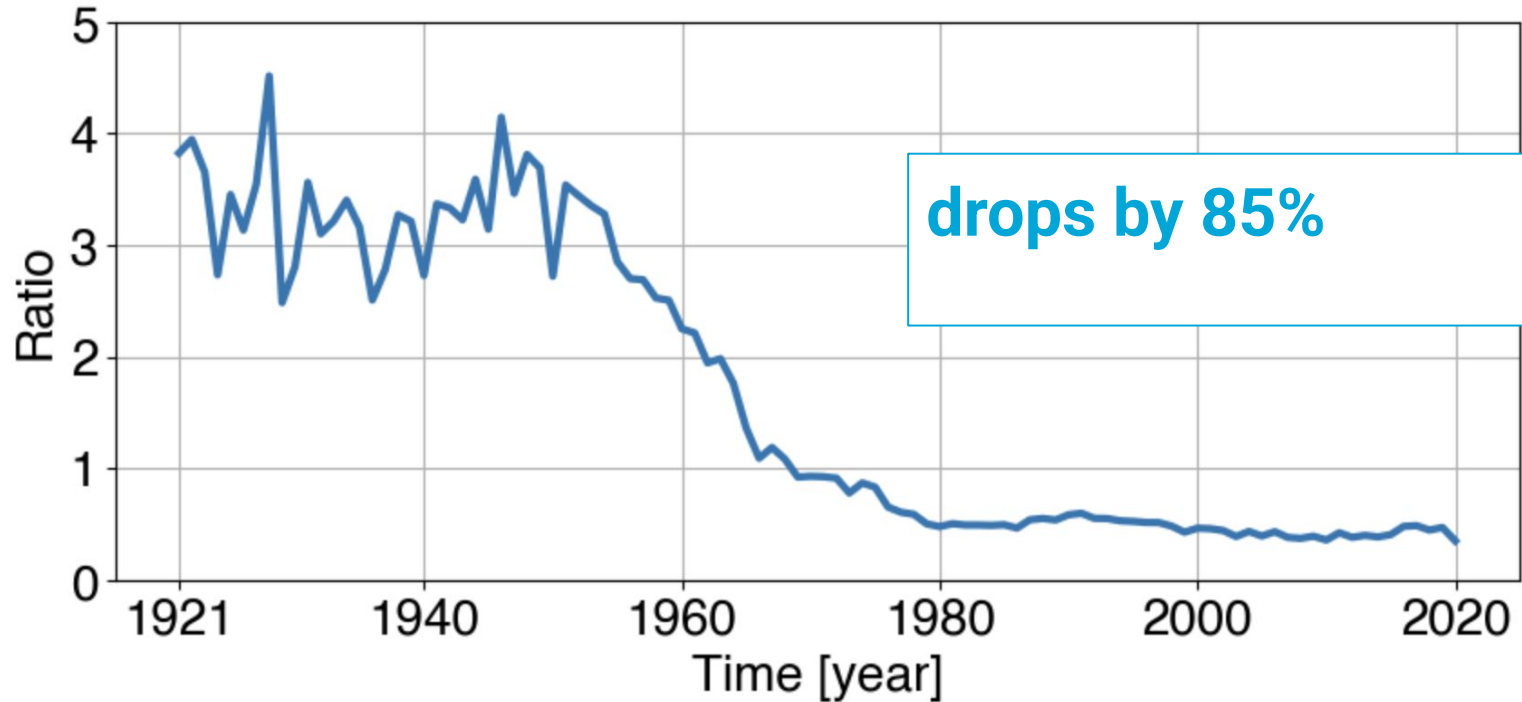
▸ Data wrangling examples

Findings (1/5) - Metrics evolution over time

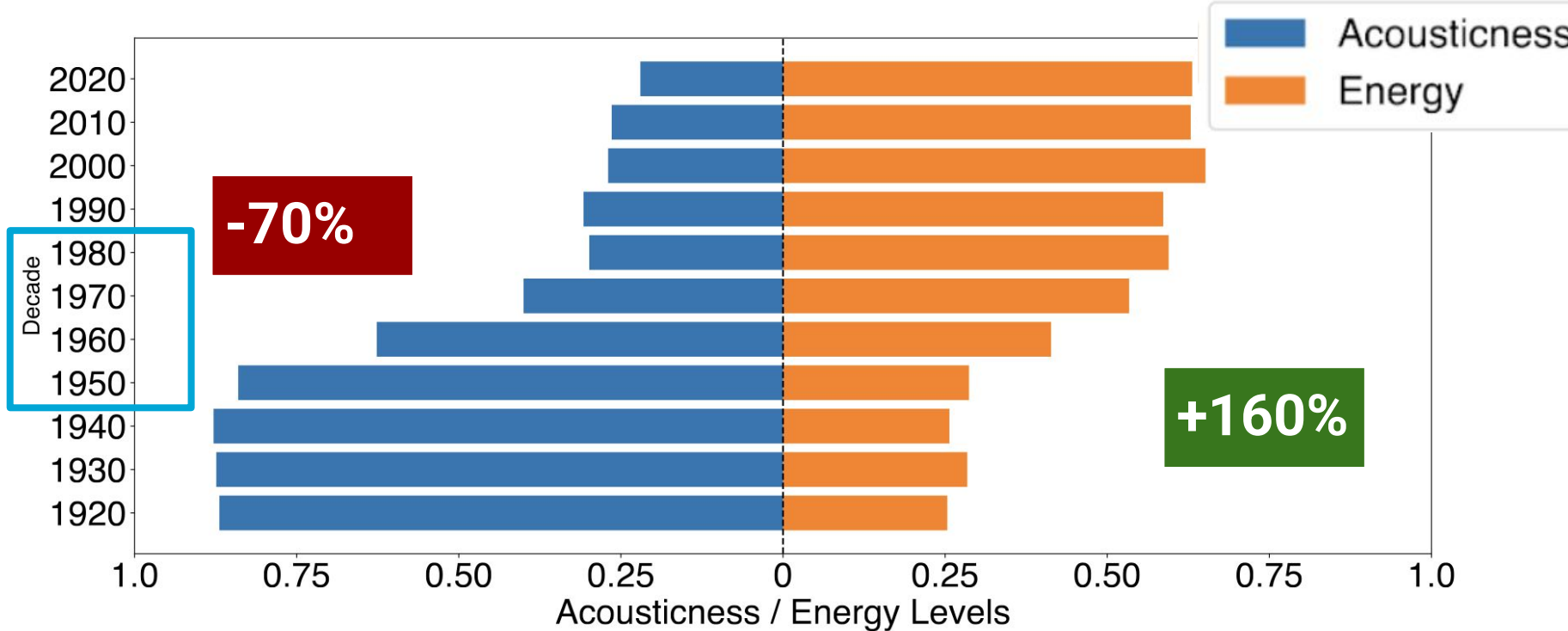


1960s-70s - key moment

Findings (2/5) - energy:acousticness evolution

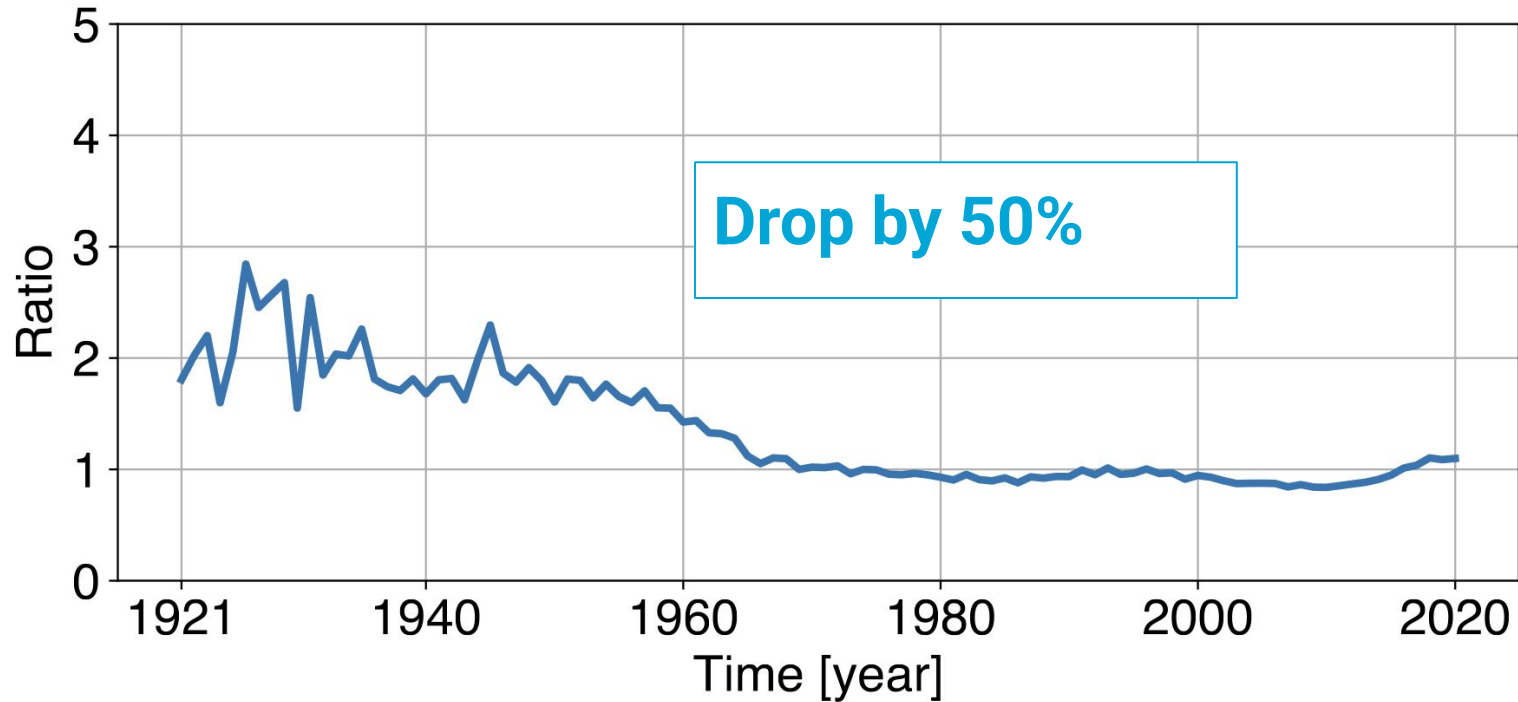


Findings (3/5) - acousticness-energy (Tukey's)

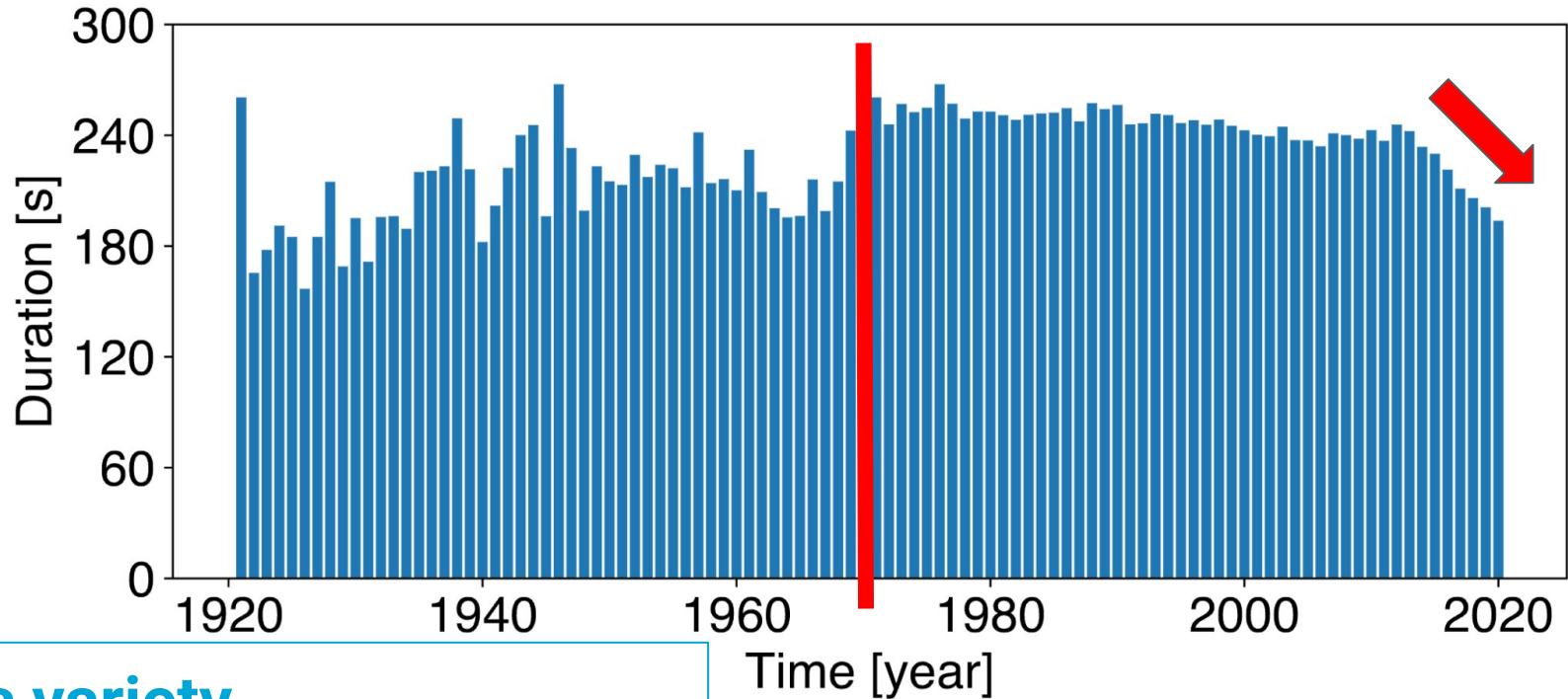


Decrease in acousticness compensated by energy

Findings (4/5) - energy:danceability evolution



Findings (5/5) - duration over time



Less variety

Decreasing in recent years

LIMITATIONS

Limitations

1. Dependence of instrumentation/production
2. Limited selection of metrics
3. Potential subjectiveness of metrics



TAKE-HOME MESSAGE

Take home message

1. Music evolution over the last century
2. MRQ: **How did music evolve over the last century?**
3. Data wrangling methods
4. Findings
5. Future



Take home message

1. Music evolution over the last century
2. MRQ: **How did music evolve over the last century?**
3. Data wrangling methods
4. Findings
5. Future

