

USE CASES

Potential real-world applications of the dataset

- Contribute to the ongoing discourse on the influence of popular music on television and the reciprocal impact of competition shows like American Idol on the music industry.

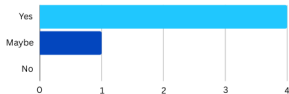
MITIGATION POSSIBLE

Maybe Yes No

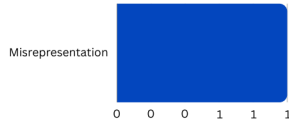
ALERT COUNT BY CATEGORY



ALERT COUNT BY MITIGATION



ALERT COUNT BY POTENTIAL HARM



American Idol

billboard
DATABASE

Created by

Winston Hermawan, Kimberly
Quiocho, and Gauri Verma

ABOUT

This dataset delves into the relationship between popular music (as defined by the Billboard Hot 100 Charts) and American Idol song performances. We joined two datasets to investigate how music popularity influences contestant song choices.

DATA CREATION RANGE:

June 2002 - May 2020

NUTRITION LABEL

Alert Count	5*
Completeness	4
Provenance	1
Misrepresentation	
Collection	0
Description	0
Composition	0

SOURCES:

https://github.com/kakey/American-Idol/blob/main/Songs/songs_all.csv
<https://github.com/HipsterVizNinja/random-data/tree/main/Music/hot-100>

The American Idol logo, featuring the words "American Idol" in a white, stylized script font inside a blue oval with a white border.

billboard DATABASE

Created by

Winston Hermawan, Kimberly
Quiocho, and Gauri Verma

5 Alerts

Week column has over-complicated date formatting

Mitigation Possible: Yes

Category: Completeness

Potential for Harm: None

Possible Mitigations: Before using week data reformat it into a more presentable format for visualizations

Week column is missing dates past a certain point

Mitigation Possible: Yes

Category: Completeness

Potential for Harm: None

Possible Mitigations: Avoid using date information past a certain point or remove incomplete rows

Data in the "isOnBillboard" column is redundant as it is implied in the "timesOnBillboard" column

Mitigation Possible: Yes

Category: Completeness

Potential for Harm: None

Possible Mitigations: Remove the "isOnBillboard" column from the dataset

Result column's data is inconsistent from time period to time period

Mitigation Possible: Yes

Category: Completeness

Potential for Harm: None

Possible Mitigations: Only use Result column's data at times when it is consistent (season to season)

Name column's data could be out of date

Mitigation Possible: Maybe

Category: Providence

Potential for Harm: Misrepresentation

- Using out of date names could be misrepresentative of contestants current gender

Possible Mitigations: Omit contestants names or double check correctness within the data set