Albums & Algorithms Course Overview and Selected Materials

(Syllabus and other resources available upon request)

Course Description:
Music is a constant, from hip-hop to country to gospel. We can play it on our phones, computers, and turntables. You have access to music from the 1700s, 1920s and 2021. In this class we will explore lyrics, beats, music history, and the health and wealth of artists and the music industry using tools from statistics and data analysis. This course helps students explore and consider potential explanations for different phenomena they might observe while learning about music, such as how hip-hop record sales have changed throughout the last 25 years. The purpose of this course is to: (a) show how statistics and data analysis are inherently creative and visual, (b) expose students to how statistics and data function in their everyday lives, (c) explore how research questions are formed, and (d) explain how data are collected/managed, analyzed, and presented visually and in written form. By exploring changes in lyrics over time we can describe how rap's language has evolved, or by looking at artists' royalties from various media we can better understand the chances of a new artist being able to survive. This course will provide a basic overview of quantitative measurement and associated quantitative concepts and will explore the ways in which certain data analytic techniques and associated quantitative models could be used to explore problems in the music industry. Finally, and most importantly, this course will help students to become more fluent in their understanding of and communication about data by moving away from data and statistics as content that is highly theoretical and moving toward content that has real-world applications.

At the end of this course, you will be able to:

- Describe and explain how research questions lead to different forms of measurement, which then lead to quantitative models that describe real world phenomena.
- Recognize the limitations of mathematical models (e.g., linear models in regression when the data suggests diminishing returns).
- Perform simple mathematical computations (i.e., means, standard deviations, correlations) in different data analysis programs (such as Google Sheets), and by hand as appropriate, associated with quantitative models, and make conclusions based on the results.
- Recognize, use, and appreciate mathematical thinking for solving problems that are a part of everyday life.
- Describe and explain the various sources of uncertainty, error, and limitations in empirical data.
- Retrieve, organize/manage, and analyze data associated with a quantitative model.
- Communicate logical arguments and their conclusions.

*Note: Each student learning outcome is reflective of the GenEd Quantitative Literacy Learning Goals, especially information literacy that emphasizes the importance of identifying reliable models and sources of knowledge.

Selected Course Resources/Readings for Integrating Music and Quantitative Literacy

The following sources provide data sets and visualizations used to guide conversations about and practice with fundamentals of quantitative analysis, including understanding empirical vs. philosophical questions, operational definitions, measures of central tendency, distribution, and data visualization, as well as analyses like ANOVA.

Daniels, M. (2017, September). The Language of Hip Hop. [Blog Post]. https://pudding.cool/2017/09/hip-hop-words/

Daniels, M. (2018, August 30). Emo rap vs. Dashboard Confessional. https://pudding.cool/2018/08/emo-rap/

Daniels, M. (2019, January 21). Rappers, sorted by the size of their vocabulary. https://pudding.cool/projects/vocabulary/

Wilber, J. (2018, June 5). The good, the bad, and the gnarly: An exploration into the music of skateboarding. Retrieved from https://pudding.cool/2018/06/skate-music/

The following video is used to facilitate an introductory conversation about how data and data visualization can be used to tell a story. Students discuss stor(ies) that grab their attention and discuss how the information used as the criteria for “best selling” changes throughout history.

Data is Beautiful. Best selling music artists 1969–2019. YouTube. https://www.youtube.com/watch?v=a3w8I8boc_I

Results

Mixed Model

Model Info

Info	
Estimate	Linear mixed model fit by REML
Call	sse_composite ~ 1 +(1 l TUID)
AIC	410.666
BIC	422.647
LogLikel.	-203.692
R-squared Marginal	0.000
R-squared Conditional	0.628
Converged	yes
Optimizer	bobyqa

[3]

Model Results

Fixed Effect Omnibus tests

F	Num df	Den df	p

Fixed Effects Parameter Estimates

			95% Confidence Interval				
Names	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	4.33	0.103	4.13	4.54	73.4	42.1	< .001

Random Components

Groups	Name	SD	Variance	ICC
TUID	(Intercept)	0.779	0.606	0.628
Residual		0.599	0.359	

Note. Number of Obs: 162 , groups: TUID 75

Mixed Model

Model Info

Info	
Estimate	Linear mixed model fit by REML
Call	sse_composite ~ 1 + wave+(1 | TUID)
AIC	376.4688
BIC	395.6573
LogLikel.	-187.6534
R-squared Marginal	0.0809
R-squared Conditional	0.7396
Converged	yes
Optimizer	bobyqa

[3]

Model Results

Fixed Effect Omnibus tests

	F	Num df	Den df	p
wave	42.9	1	94.2	< .001

Note. Satterthwaite method for degrees of freedom

Fixed Effects Parameter Estimates

			95% Confidence Interval				
Names	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	3.685	0.1406	3.409	3.960	157.0	26.20	< .001
wave	0.341	0.0521	0.239	0.443	94.2	6.55	< .001

Random Components

Groups	Name	SD	Variance	ICC
TUID	(Intercept)	0.793	0.630	0.717
Residual		0.499	0.249	

Note. Number of Obs: 162 , groups: TUID 75

Mixed Model

Model Info

Info	
Estimate	Linear mixed model fit by REML
Call	sse_composite ~ 1 + wave + female + non_white + fg_dummy+(1 | TUID)
AIC	328.841
BIC	360.887
LogLikel.	-162.930
R-squared Marginal	0.141
R-squared Conditional	0.781
Converged	yes
Optimizer	bobyqa

[3]

Model Results

Fixed Effect Omnibus tests

	F	Num df	Den df	p
wave	43.9792	1	88.7	<.001
female	0.0428	1	67.6	0.837
non_white	5.8715	1	70.3	0.018
fg_dummy	4.34e-5	1	126.1	0.995

Note. Satterthwaite method for degrees of freedom

Fixed Effects Parameter Estimates

			95% Confidence Interval				
Names	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	3.95520	0.2049	3.554	4.3567	97.8	19.30692	<.001
wave	0.33621	0.0507	0.237	0.4356	88.7	6.63168	<.001
female	-0.04320	0.2087	-0.452	0.3659	67.6	-0.20700	0.837
non_white	-0.51121	0.2110	-0.925	-0.0977	70.3	-2.42312	0.018
fg_dummy	0.00134	0.2029	-0.396	0.3990	126.1	0.00658	0.995

Random Components

Groups	Name	SD	Variance	ICC
TUID	(Intercept)	0.771	0.595	0.745
Residual		0.451	0.204	

Note. Number of Obs: 149 , groups: TUID 70

Mixed Model

Model Info

Info	
Estimate	Linear mixed model fit by REML
Call	sse_composite ~ 1 + wave + female + non_white + fg_dummy + evt_composite_wcost+(1 I TUID)
AIC	307.129
BIC	346.844
LogLikel.	-153.433
R-squared Marginal	0.298
R-squared Conditional	0.748
Converged	yes
Optimizer	bobyqa

[3]

Model Results

Fixed Effect Omnibus tests

	F	Num df	Den df	p
wave	57.1925	1	96.5	<.001
female	1.86e-4	1	62.9	0.989
non_white	7.3840	1	65.6	0.008
fg_dummy	0.0551	1	109.7	0.815
evt_composite_wcost	24.5771	1	132.8	<.001

Note. Satterthwaite method for degrees of freedom

Fixed Effects Parameter Estimates

			95% Confidence Interval				
Names	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	2.27703	0.3889	1.515	3.039	128.8	5.8551	<.001
wave	0.41282	0.0546	0.306	0.520	96.5	7.5626	<.001
female	0.00239	0.1751	-0.341	0.346	62.9	0.0137	0.989
non_white	-0.48251	0.1776	-0.831	-0.134	65.6	-2.7174	0.008
fg_dummy	0.04270	0.1819	-0.314	0.399	109.7	0.2348	0.815
evt_composite_wcost	0.35803	0.0722	0.216	0.500	132.8	4.9575	<.001

Random Components

Groups	Name	SD	Variance	ICC
TUID	(Intercept)	0.617	0.380	0.641
Residual		0.462	0.213	

Note. Number of Obs: 148 , groups: TUID 70

Mixed Model

Model Info

Info	
Estimate	Linear mixed model fit by REML
Call	sse_composite ~ 1 + wave + female + non_white + fg_dummy + belong_class_composite + belong_faculty_composite + belong_student_composite + evt_composite_wcost+(1 | TUID)
AIC	304.308
BIC	365.062
LogLikel.	-155.046
R-squared Marginal	0.360
R-squared Conditional	0.753
Converged	yes
Optimizer	bobyqa

[3]

Model Results

Fixed Effect Omnibus tests

	F	Num df	Den df	p
wave	33.2952	1	105.9	<.001
female	0.1005	1	58.4	0.752
non_white	7.1461	1	57.0	0.010
fg_dummy	0.4059	1	107.2	0.525
belong_class_composite	0.0911	1	131.6	0.763
belong_faculty_composite	6.6758	1	138.8	0.011
belong_student_composite	0.7678	1	137.8	0.382
evt_composite_wcost	15.2414	1	129.6	<.001

Note. Satterthwaite method for degrees of freedom

Fixed Effects Parameter Estimates

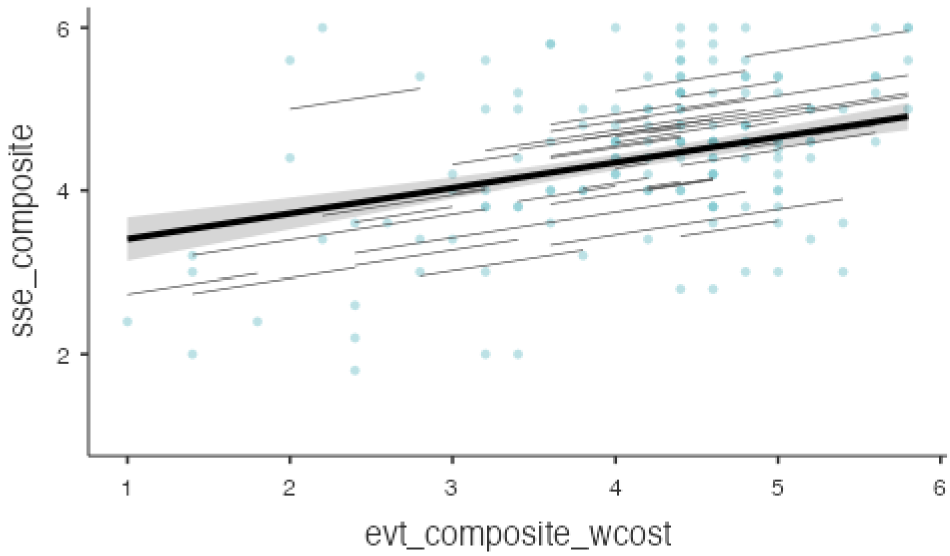
			95% Confidence Interval				
Names	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	1.4335	0.4735	0.5055	2.362	105.1	3.028	0.003
wave	0.3564	0.0618	0.2353	0.477	105.9	5.770	<.001
female	-0.0542	0.1708	-0.3889	0.281	58.4	-0.317	0.752
non_white	-0.4580	0.1713	-0.7938	-0.122	57.0	-2.673	0.010
fg_dummy	0.1150	0.1805	-0.2388	0.469	107.2	0.637	0.525
belong_class_composite	0.0179	0.0595	-0.0986	0.134	131.6	0.302	0.763
belong_faculty_composite	0.1788	0.0692	0.0432	0.314	138.8	2.584	0.011
belong_student_composite	0.0461	0.0526	-0.0570	0.149	137.8	0.876	0.382
evt_composite_wcost	0.3136	0.0803	0.1561	0.471	129.6	3.904	<.001

Random Components

Groups	Name	SD	Variance	ICC
TUID	(Intercept)	0.583	0.340	0.614
Residual		0.463	0.214	

Note. Number of Obs: 148 , groups: TUID 70

Effects Plots



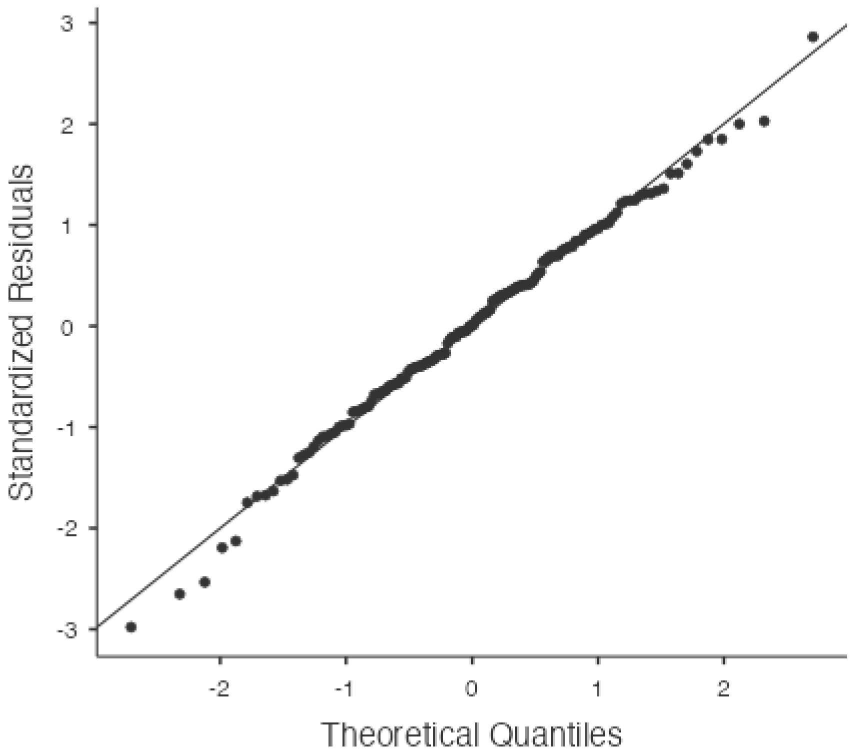
Note: Random effects are plotted by TUID

Assumption Checks

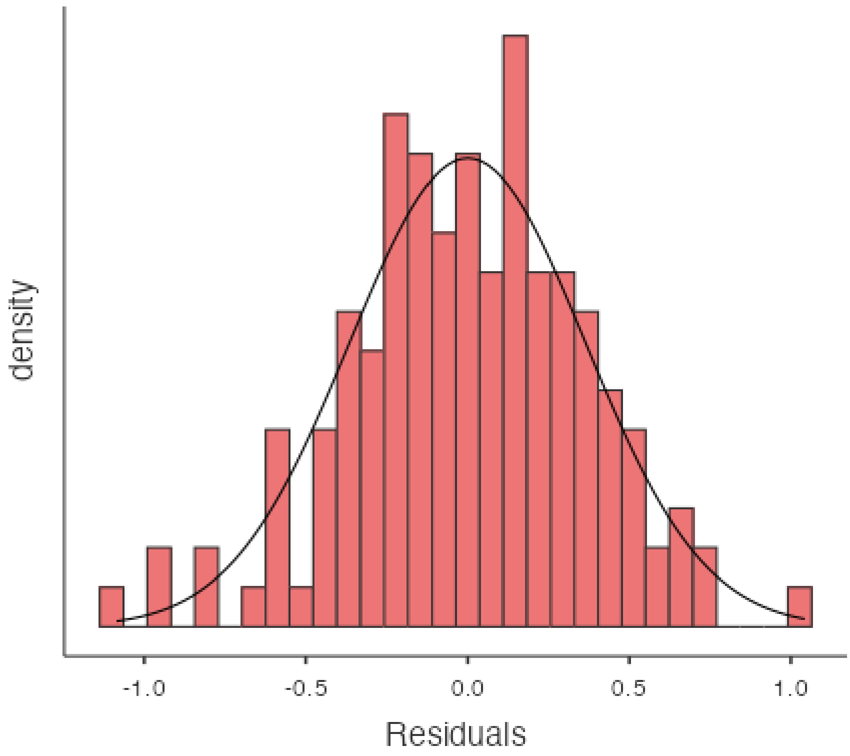
Test for Normality of residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0367	0.988
Shapiro-Wilk	0.9941	0.814

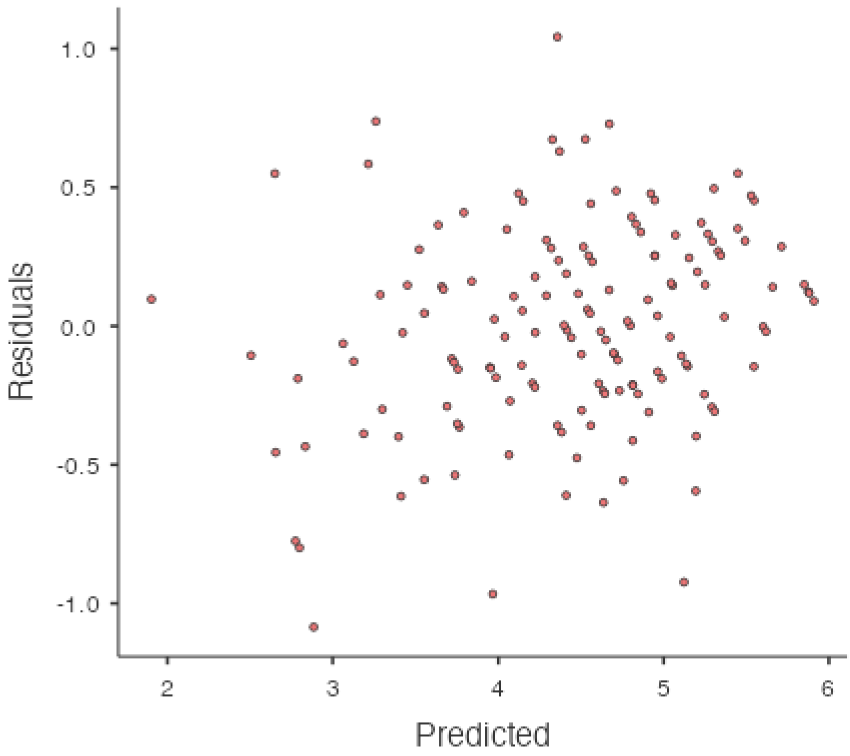
Q-Q Plot



Residual histogram



Residual-Predicted Scatterplot



References

[1] The jamovi project (2022). *jamovi*. (Version 2.3) [Computer Software]. Retrieved from https://www.jamovi.org.

[2] R Core Team (2021). *R: A Language and environment for statistical computing*. (Version 4.1) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from MRAN snapshot 2022-01-01).

[3] Gallucci, M. (2019). *GAMLj: General analyses for linear models*. [jamovi module]. Retrieved from https://gamlj.github.io/.

Qualitative Analysis of Open-Ended Responses

Q1: What aspects of the course are supporting your engagement and learning? How?

Code	Time 1	Time 2	Time 3	Total
Course Structure	9	8	13	30
CS - Lecture	6	4	2	12
CS - Recitation	6	7	12	25
Pedagogy	11	21	13	45
P - Repetition/Review	2	4	2	8
P - Discussion/Participation	5	5	2	12
P - Collaboration	4	14	9	27
Course Content	33	22	13	68
CC - Music	23	14	5	42
CC - Stats	5	1	1	7
CC - Labs	4	3	4	11
CC - Homework	3	0	2	5
CC - Project	1	4	2	7
CC - Materials	5	5	2	12
Instructors	20	12	16	48
I - Professor	14	8	10	32
I - TAs	13	10	12	35
Support	8	9	13	30
S - Peer Support	2	5	5	12
S - Faculty Support	6	4	8	18
Atmosphere/Environment	6	7	2	15
I Don't Know/Nothing	1	0	2	3
Blank	1	3	7	11
Total Code Applications	188	170	157	515

Time 1:

Course content being tied to music was perceived as supportive of student engagement and learning.

Both the professor and TAs were perceived as supportive of engagement and learning.

The course's accessible faculty, emphasis on discussion and participation, and inclusive environment were perceived as supportive.

> *The music aspect of the class is absolutely keeping me interested and supporting my learning. It is one of the most important things to me and I always enjoy working with it. It works as a balance for me with this class that almost cancels out the actual work.*

> *My TA is very enthusiastic and is willing to answer students' individual questions. My professor makes the course feel more relax and inclusive instead of just throwing numbers and terms at me*

group work and open discussion because it makes me feel more comfortable sharing or asking if i don't understand something

Time 2:

Course content being tied to music continued to be a frequently referenced supportive element of the course.

The pedagogical approaches used in the course, especially the collaborative in-class work and group project, emerged as a supportive element of the course.

Instructors continued to be perceived as supportive, but direct mentions were less frequent.

> *Having the PowerPoints to review with if I missed something during the lecture. Being able to talk with my TA and professor if I don't understand something or just need clarification. They give very good explanations and examples. Having similar work done in class/lecture and the music (DJ breaks) are helpful and fun to make class interesting and easier to understand and look back on when we apply it to [our] own work later.*

Time 3:

Instructors were perceived as supportive of engagement and learning, and faculty support (i.e., help-seeking during/outside of class) emerged as a more important area of support.

Recitation was perceived as supportive of engagement and learning.

Collaboration was perceived as supportive.

Music was less emphasized as supportive.

> *Having a TA to go over material and answer question not only in person but also through email. Very helpful*

> *Recitation is helpful because it is a smaller setting and I feel more comfortable asking questions and working with others.*

> *The group work for labs every week helps me form relationships with students easier. Because of this, I feel like I am able to do better on my assignments.*

Q2: What aspects of the course are challenging your engagement and learning? How?

Code	Time 1	Time 2	Time 3	Total
Course Content	28	19	21	68
CC - Math/Statistics	20	13	15	48
CC - Technology/Project Design	9	4	1	14
CC - GenEd	1	1	0	2
Course Structure	13	11	10	34
CS - Lecture	7	5	5	17
CS - Recitation	1	0	1	2
CS - Time	6	6	4	16
Course Work	11	17	12	40
CW - Homework	2	4	2	8
CW - Labs	3	7	2	12
CW - Quizzes	0	2	1	3
CW - Projects	3	2	2	7
CW - Workload	0	2	0	2
CW - Group Work	4	3	0	7
CW - Availability	1	3	3	7
Student Challenges	12	11	4	27
SC - Disinterest	4	5	1	10
SC - Learning Styles	5	1	0	6
SC - Other Factors	3	5	3	11
Whole Course	1	0	2	3
No Challenge	10	7	4	21
Positive	7	9	4	20
NA/Blank	1	3	9	13
Total	152	140	106	398

Time 1:

Course content, especially math and statistics and to a lesser extent using Excel and other analytic tools, was perceived as challenging.

Course structure, in particular the large lecture, class length, and meeting time, were perceived as challenging.

Students described some personal challenges such as learning differences, disinterest, and difficulty focusing.

Some course work was perceived as challenging, but emphasis was not placed on a particular type of course work.

A number of students felt the course was not challenging, felt the challenge was appropriate, or expressed approval of the course.

The course is challenging to me because I've never taken a stats class and have no experience using google sheets. That being said, I feel like I've already learned a great amount only three weeks into the class.

the whole idea of statics and graphs scare me, especially excel. Just because 2 semesters ago I was in a research class and all we did was excel and it was confusing, but maybe that is because he explained it horribly as a teacher because he already understood it. I am just nervous about that part honestly.

The setting of a lecture hall is harder to pay attention in, for me at least.

Time 2:

Fewer students identified math/statistics as a challenge.

Course work received greater emphasis, with students finding labs, and to a lesser extent homeworks, challenging or overly complicated.

Students continued to describe personal challenges that impacted their course experience.

A number of students felt the course was not challenging or described positive feelings about the course.

Course structure, especially the lecture format and class length and meeting time, continued to present a challenge for some students.

I find the lectures a little hard sometime because I have a hard time sitting still and listening, but the breaks always help me recoup.

I think sometimes the labs and hw feel a little bit tedious but over all I don't think it really is a challenge engaging and learning.

Sometimes the content (wording or framing of a question) on homework and quizzes can be a little bit confusing or unclear

Time 3:

The number/percentage of respondents describing math/statistics as a challenge from T2 to T3 increased.

Course work and structure continued to be identified as challenges for some students.

math is still a challenge, but it is getting easier

Numbers terrify me. Statistics and understanding what they represent is not easy for me. I get lost and confused easily.

The material because it is difficult and causes me stress when doing hw

i have trouble sitting through lectures so that is a bit of a struggle for me to listen and try to apply what i've learned. this class is structured better than other classes though, and i like getting breaks or when things are interactive