

Final Team Project

Final Project deliverables		Due date
1	Project executive summary	11:59pm (end of the day), October 11, 2020
2	Dashboard	11:59pm (end of the day), October 11, 2020
3	Appendix/Supplementary materials	11:59pm (end of the day), October 11, 2020

Final Project Deliverables:

1. **Executive summary:** summarize your findings for the board of directors of Universal Music. Since they are not all very well-versed in excel and modeling techniques, you need to formulate and communicate actionable business recommendations based upon your analysis, including its limitations, in easy-to-understand terms. Present your reasoned opinion on the efficacy and viability of your analytics-based recommendations for Universal Music and critically assess their validity. If you feel your conclusions and recommendations are not sufficiently supported by data, include your recommendations for improving them.

Executive summary preparation guidelines: should not exceed more than 4 pages

2. **Dashboard:** present key findings in a dashboard.

Dashboard preparation guidelines: your key findings should be presented on one page.

Final Project Objectives:

You are expected to demonstrate your knowledge and ability to

- I. **Explain** in your own words the key ideas behind fundamental techniques in data analytics, including dashboarding, classification, clustering and AB-testing;
- II. **Identify** new opportunities to use these techniques across business domains to guide decision making;
- III. Confidently **apply** these techniques to novel problems using a combination of Excel and JMP;
- IV. **Formulate** and **communicate** actionable business recommendations based upon your analysis, including its limitations;
- V. Critically **assess** the validity of analytics-based recommendations in the context of specific business decisions.

Final Project Submission Instructions:

- I. Final team project deliverables must be submitted **via blackboard** by **only one team member by the due date indicated in the table above**.
- II. One team member submits the following items:
 - a. Zipped folder that contains the following deliverables by **by 11:59pm (end of the day), October 11, 2019:**
 - i. **Project Executive Summary** (up to 4 pages);
 - ii. **Dashboard** (1 page word file, **THERE IS NO TEMPLATE**);
 - iii. **Appendix** (not mandatory, only if you want to include additional information, results that are not included in the executive summary and dashboard, assumptions, etc. No page limit for the appendix. **THERE IS NO TEMPLATE**);
 - iv. **Excel file** with analysis. Make sure your Excel file contain analyses that justify what is included in the executive summary and dashboard. No limit on the number of sheets in the excel file.
- III. You may make multiple submission. Only the last submission will be grade.

Tips for developing an effective team:

- Do team members want to participate in the team? Do team members attend all meetings?
- Are members committed to accomplishing the team mission and expected outcomes?
- Has the team defined and communicated its goals; its anticipated outcomes and contributions; its timelines; and how it will measure both the outcomes of its work and the process the team followed to accomplish their task.
- Do team members hold each other accountable for project timelines, commitments, and results?
- Do all team members understand the roles and responsibilities of team members?
- Are the team members clear about the priority of their tasks?
- Do team members communicate clearly and honestly with each other?

Case: Music in the 2010's

An Overview of the Music Industry

The way in which people listen to music has changed throughout the years. From the cassette players of the 1970s, to the extreme market disruption brought about by iTunes in the early 2000s the way in which artists reach their fans has changed rapidly. The 2010s saw another massive disruption to the music industry: the rise of streaming services, like Spotify. These services forgo the need for listeners to own music altogether, and instead deliver large libraries of music to their subscribers for a monthly fee. They also allow artists to interact more directly with their fans, and understand their audiences through analytics.

Speaking of analytics, let's discuss Spotify's data assets. Spotify does a lot more than simply act as an intermediary between artists and fans. They also gather immense amounts of data on the music on their platform, and the artists that create it. For example, Spotify's algorithms can identify: the key of a song, its mode (major or minor), and its genre (rap, pop, country, etc.). They also attempt to quantify more abstract characteristics of songs, like how happy a song is, how easy it is to dance to it, etc. At a more aggregate level, Spotify tracks popular songs, which ultimately find their way into popular playlists, and if a song is particularly good it will find its way into a "top 10" playlist.

Because Spotify is a now major player in the music industry, music companies around the world endeavor to understand what is referred to as "the Spotify ecosystem." They want to know how to best promote their music on Spotify, how to best support their artists, how to get their songs recommended, and of course how to make a name for themselves. In many ways, this is incredibly different from how the music industry used to operate. The Spotify ecosystem is highly dynamic, representing peoples' tastes, and is very susceptible to shocks. For example, when an artist releases an instant hit you can bet the dynamics of this ecosystem will change because of it.

For a music company to capitalize on a market like this, it must also be dynamic. It must know the real-time preferences of listeners, the state of the ecosystem, and have an idea of where it is going. To grow any artist, this company needs to know how to best support them, from recommending up-and-coming artists to listeners to utilizing its well-made artists and providing them the support they need.

Monetize Spotify

Monetizing Spotify Spotify claims that it pays \$0.006-\$0.0084 to the rights holders for a song every time that song is streamed. While this might seem like a small amount of money, when you consider a song like P!nk's "Walk me Home" has been streamed roughly 208,500,000 times, that translates into anywhere between \$251000 and \$1751400 for that one song alone. Considering P!nk is a very big name in the music industry, that revenue is probably closer to the upper end of that range, and possibly even higher. This leads us to an interesting question: how do we convert the popularity of a song to revenue? There are two options for Universal Music to contract with Spotify. Note that in either of these options each stream generates the same amount of revenue for that song.

Option #1:

The following scheme is a very realistic way of modeling this. Suppose Universal Music is about to release a new Taylor Swift album, and they plan to release it on Spotify. As Taylor Swift is one of their premium artists and this is a new album, they are going to require Spotify pay them \$.0084 every time a song from that album is streamed. If the album truly is as popular as Universal Music bet it would be, everyone is happy. However, what happens if the album is an unexpected flop? In that case, Spotify's contract with Universal Music specifies that rather than paying \$.0084 per stream, they now only pay Universal Music \$.002 per stream. This allows Spotify to reallocate money towards more popular artists on their platform, while allowing Universal Music to recoup revenue.

Option #2:

Another option for Universal Music is to charge a flat fee of \$.006 per stream, regardless of the song being actually popular or not. Of course, if Universal Music plays it safe and only requires Spotify give them \$.006 per stream for a given song that ends up being a huge hit, there is an opportunity cost on the part of Universal Music. Universal Music needs to choose the one of the two options to contract for a song before the song is being released on Spotify. Universal Music's goal, then, should be able to accurately predict the popularity of a song. That way they can make a premium off of songs that are really popular, while carefully assessing the true value of their music.

Assignment

- Despite the important role data now plays in the music industry, many companies are behind the “data curve.” To that end, Universal Music has hired you as a data analyst to help them understand how they can utilize Spotify to their benefit. The company has data on a large number of songs on Spotify. It is the hope of Universal Music that they can gain an understanding of trends in what characteristics popular songs have, they can better utilize their artists. Below are some of the questions Universal Music would like answered: Do you see any trends for this industry? Any high-level directions that Universal Music should move towards?
- Are there any factors that lead to the success of a popular song? Are the popular songs different in terms of the variables measured by Spotify or the KPIs you came up with? If not, what might this mean in terms of what makes a song popular?
- Is there a strategy to maximize the expected gain given the contract options?

Data Description

The data for this project contains characteristics on 88386 songs on Spotify. The characteristics are defined as follows:

- **Genre:** The genre of the song (rap, rock, country, etc.).
- **Artist_name:** The name of the artist of the song.
- **Track_name:** The name of the song.
- **Track_id:** A unique id for the song.
- **Popularity:** Whether or not a track is considered “popular.” This variable depends on factors such as: how often the song has been listened to, and how recently it has been listened to. (1 for popular and 0 otherwise)
- **Acousticness:** A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
- **Danceability:** How suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.
- **Duration_ms:** The duration of the track, in milliseconds.
- **Energy:** A measure from 0.0 to 1.0 that represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.
- **Instrumentalness:** Detects whether a track contains no vocals. “Ooh” and “aah” sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly “vocal”. The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0
- **Key:** The key of the song
- **Liveness:** Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live
- **Loudness:** The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typical range between -60 and 0 db.
- **Mode:** Whether the song is in a major or minor key
- **Speechiness:** Detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.
- **Tempo:** The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.
- **Valence:** A measure from 0.0 to 1.0 describing the *musical* positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).
- **Streams:** The number of streams a song has received. **Note: This variable cannot be used directly for modeling purposes, only for determining revenue after popularity has been determined. (In other words, once the popularity of the song is revealed, revenue can be calculated based on streams and the fee stated in the contract chosen.)**

General Business Analytics Project Guidelines

Project Phases:

- **Business Understanding**
 - Determine business goals
 - a business goal states objectives in business terminology
 - Determine data analysis goals
 - a data analysis goal states project objectives in technical terms
 - Create hypotheses and indicate assumptions
- **Data Understanding**
 - Explore the data
 - Verify the quality
 - Find outliers

Start with initial data collection and proceed with activities in order to get familiar with the data. Identify data quality problems, find insights, and form hypotheses.

- **Data Preparation**
 - Construct data
 - includes constructive data preparation operations such as the production of derived attributes, entire new records, or transformed values for existing attributes
 - Integrate data
 - includes methods whereby information is combined from multiple tables or records to create new records or values
 - Format data
 - modifications made to the data that do not change its meaning but might be required by the model or modeling tool
- **Modeling**
 - Select the modeling technique
 - include rationale for the selection
 - Build model
 - include parameter settings
 - Assess model
 - rank the models, where necessary

Various modeling techniques should be selected and applied, and their parameters should be calibrated to optimal values (if possible). Some techniques have specific requirements on the form of data. Therefore, stepping back to the data preparation phase is often necessary (for example, converting a numeric variable to a binary variable for the classification decision tree).

- **Evaluation**
 - Evaluate the model
 - how well did it perform on test data?
 - Interpretation of model

Thoroughly evaluate the model and review the steps executed to construct it; be certain that it properly achieves the business objectives. At the end of this phase, a decision on the use of the data mining results should be reached.

- **Deployment**

- Determine how the results need to be utilized
 - who needs to use them?
 - how often do they need to be used?
- Deploy results
 - describe the deployment process, or outline the deployment plan

- **Recommendations**

- Formulate and communicate actionable business recommendations based upon your analysis
 - Include their limitations
 - how often do they need to be used?
- Critically assess the validity of analytics-based recommendations in the context of specific business decisions

The knowledge gained will need to be organized and presented in a way that is consumable by the customer. However, depending on the requirements, the deployment phase may be as simple as generating a report or as complex as implementing a repeatable data analysis process across the enterprise.

Grading:

Overall, project grades will be assigned based on the executive summary and a dashboard. Students will be assessed not only on the accuracy of the analysis, but its quality and depth, clarity, validity and viability of their data derived business recommendations, the analytics tools used, clearly communicated evidence supporting their recommendations. Detailed grading rubrics are posted on blackboard.

Criterion	Needs Improvement (1)	At Expectations (3)	Exemplary (5)
Clarity	Approach and analysis are not clearly described.	Approach and analysis clearly and concisely described.	Approach and analysis clearly and concisely described, including why these were the most effective options in this particular business context
Accuracy	Techniques are applied incorrectly, or interpretations are not supported by data	Techniques are applied correctly, with interpretation of results	Techniques are applied correctly with in-depth, high-quality analysis connected to the business context
Organization	Flow is difficult to follow	Flows smoothly and follows standard business writing conventions	Report is logically ordered to convey insight and argue for business recommendation. Visualizations support main message.
Business Acumen	No or very minimal business acumen built in the project.	A team demonstrated an ability to deliver some tangible business value.	A team demonstrated an insightful view of the overall business landscape, including financial and customer/client business drivers, and applied that understanding or interpretation to the project.
Supporting evidence	Recommendations are only loosely related to analysis and evidence	Recommendations strongly motivated by analysis and evidence	Recommendations strongly motivated by analysis and evidence, complete with possible limitations
Enterprise Perspective	Solution is not clear, or is too complicated to be implemented in an enterprise environment; Not going to work.	Solution is clear and can be implemented in an enterprise environment pending modifications.	The presenter integrated business and technology requirements to present recommendations that could work across the enterprise.
Business Value	Solution does not appear to improve business processes or add any value to the business. No business validation is presented.	Solution moderately improves business process and provides positive impact to the business.	Solution greatly improves business process and provides substantial positive impact to the business. Solution may also be leveraged for other similar business problems.
Creativity	Follows directions without understanding the purpose for innovation, does not offer new ideas.	Offers new ideas, but stays within narrow perspectives.	The team demonstrated creative thinking and problem solving in their approach. They were not afraid to explore new opportunities or techniques that could add business value. Used ingenuity and imagination, going outside conventional boundaries.
Analytical/Conceptual/ Critical Thinking	Little or no evidence of critical, analytical, conceptual thinking skills.	Demonstrated evidence of critical, analytical, conceptual thinking skills	Demonstrated strong ability and competence to identify key challenges and select effective solutions by skillfully conceptualizing, applying, analyzing, synthesizing, evaluating and clearly communicating information and results.
Other Analysis	Analysis and recommendation are largely limited to one very specific basic business question.	Some insights for the business beyond the primary/initial/stated business question are discussed	High-quality insights and recommendations are presented that extend beyond primary/initial/stated business question, illustrating creative uses of the dataset.