Slide 1

Good evening everyone and thank you for attending our presentation. We have spent the semester working on Project Venue Scraper. This project began as a prototype I created prior to this semester to aid my job as a booking agent for bands.

Slide 2

My name is CJ, and I served as the project manager. James served as our Database Designer, Katelyn as both our Requirements Analyst and UX Designer. Raegan was our team’s Business Analyst AND Researcher, while Sha and Fatima both aided as our Software Architects.

Slide 3

Going into this project we learned that all of us have conflicting work schedules which makes scheduling meetings difficult. Making flexibility worse, my day job occasionally has me traveling out of state for a week with short notice. To circumvent scheduling issues, we decided to have Microsoft Teams as our primary channel of communication. Teams provided the mobility and flexibility necessary for communicating when Laptop / Desktop access was impossible. Teams also served as our group’s repository as we exchanged files, with the added bonus of viewing edits live as we worked together.

This group took an Agile approach to our project, with the caveat of omitting scheduled weekly meetings due to the previously mentioned scheduling conflicts. Instead, we used Teams as a bulletin board for efficient digital communication, with each week’s task assignments existing as its own post on our feed.

This course was structured into weekly sprints where each of us had individual tasks serving the greater goal. Once individual work was completed our team came together to combine our elements into various reports and presentations leading up to tonight. As our team’s repository would reflect, we stayed organized by maintaining separate folders for each Sprint, along with each

We had the benefit of my expertise of what Venue Scraper needs to do and through multiple discussions with my teammates we were able to craft user stories for each type of user who I intend on using this software.

Slide 4

What is the Venue Scraper project exactly? This software was designed for Shirley Road Records, a Durham-based music business, to maximize projected ticket sales. The prototype originally could only manage manually inserted data for concerts. This process is slow and needs the benefit of automation to scale up for use as a forecasting model. This is where the name Venue Scraper originates, as I had made a web scraper to deliver data on individual concerts from a venue in my market. I was pleased with the outcome and knew creating multiple scrapers for multiple venues was a way to speed up the data collecting and cleaning. By performing a ridge regression on the data, what outputs is a loosely ranked list of which artists, venues, ticket prices, and days of the week maximize concert attendance. The core of the project is to answer the question “how do I minimize my odds of losing money booking a concert.”

Slides 5 - 10

TURN IT OVER TO MY TEAMMATES TO EXPAND ON WHAT I’VE LAID OUT.

Slide 11: James

Slide 12: Security

Security for both our resources, and our userbase, is a concern for any software developer. Our software would follow similar security guidelines to what are popular today. Two-factor authentication serves as a significant barrier for hackers and other nefarious actors to overcome. We intend to integrate this feature well before this reaches users.

Multi-layer encryption for the database and user activity will aid in ensuring that if there is a security breach, our data and all user data will be safe. We know that if there is fraudulent activity within the database, such as deleting concert information or injecting fake records, the math behind the forecasting loses all integrity, and thus erodes usability and trust for our userbase.

Finally, we will maintain logs of all activity with the software. By keeping track of the actions of our users, APIs, and web crawlers, we are empowered to identify any threats, intentional or not, and enhance security measures.

Slide 13: Katelyn

Slide 14: Architecture (and James covers Database)

The project operates by communication between components designated to specific tasks. We have three distinct components at play: data collection, data storage, and data analysis. Through dissecting our prototype, we have found methods for each component to receive maintenance without impacting other components. Some examples are the ability to modify web scrapers independent from Storage and Analysis, expanding our database with no affect to our Data Collection, and the ability to test different statistical methods without interacting with our Collection and Storage activity.

As it currently exists with the prototype, data collection happens in a Python environment, Excel is used for storage, and regression analytics occur within R.

Slide 15: Raegan

Slide 16: Sha

Slide 17: James

Slide 18:

As I previously mentioned, our main challenge was navigation around everyone’s conflicting schedules. Thankfully, by using our resources, we had a (my-newt) number of instances where one person’s schedule affected another teammate’s ability to begin their tasks. I am proud to share we have taken my initial prototype, dissected it, and created a plan for future evolution into a tool that will not only go beyond my business wants and needs, but for potential use to a broader market of my industry peers.

Slide 19:

In the future we can see this software providing a ticket purchasing service. There is the potential for this project to be a web accessible SaaS offering even ticket purchasers can use. By having menus allowing for filtering by genre and location, any music fan could visit our website and narrow down upcoming concerts relevant to their taste and place.

This plays into future configuration options for our end user. We want our customers, to an extent, to have choice in their experience. As a customer myself I like it when I can switch a screen into dark mode. I like when I am given the option to have a non-English website translated to my native language. These configurations have become near mandatory for services to be considered “good quality” by today’s standards and our users deserve that experience.

Finally, while the statistics operate on existing data, the database more closely resembles a “black box” than an accessible archive. As a music nerd, I love being able to look through archival records to see where, when, and with whom my favorite bands have performed. In the future it would be nice to have the database as an accessible archive for our users to explore artists in the same way I do. As a fun aside, it would be neat to create scenarios for users to discover a new act just by seeing they had shared the stage in a tiny dive bar once upon a time before the band got big. We can provide the tools for this to become a reality.

Slide 20:

Thank you for being with us tonight. We welcome any questions you may have.