

NWA Data Science Meetup

Let's ring in the new year with a DOUBLE dose of data science!

The first meetup of 2016 will feature two exciting talks, one on big data and NoSQL technologies and the other on machine learning and sporting event prediction.

#nwadatascience - <http://www.meetup.com/NWA-Data-Science-Meetup/>

Date and Time: January 20th, 2016 @ 7pm-9pm

Location: Lofty Labs, 1 E Center Street #320-C, Fayetteville, AR 72701 (use the main entrance off the square on East Avenue)

Drinks, Snacks, & Swag provided by our sponsor MapR Technologies



Speakers and Topics:

Jim Scott, Director, Enterprise Strategy and Architecture for MapR Technologies will present on NoSQL technologies: For the past 25 years applications have been getting built using an RDBMS with a predefined schema which forces data to conform with a schema on-write. Many people still think that they must use an RDBMS for applications even though records in their datasets have no relation to one another. Additionally, those databases are optimized for transactional use, and data must be exported for analytics purposes. NoSQL technologies have turned that model on its side to deliver groundbreaking performance improvements. I will walk through a music database with over 100 tables in the schema and show how to convert that model over for use with a

NoSQL database. I will show how to handle creating, updating and deleting records, using column families for different types of data (and why). I will then show how to use the exact same data without moving or transforming it to perform analytics, by leveraging Apache Drill's ANSI-SQL capabilities on the NoSQL database.

Mark Conway, TCS data scientist and enthusiastic kaggler, will be presenting his work on statistical model development for sporting event prediction: I present a sports model to generate game predictions using a machine learning pipeline. I will examine various models using different sources of sports data to examine the intersection of data science and sports. I will also discuss an annual Kaggle competition for NCAA's March Madness with source code and visualizations.

Big thanks to Lofty Labs for hosting this meetup!

