

# Introduction

- Dave Langer, VP of Data Science – Data Science Dojo
- 20+ years in technology:
  - Roles in development, architecture, & BI/DW/analytics.
  - Last job – Sr. Director, BI & Analytics @ Microsoft.
  - Text Analytics is one of passions.
- Joined DSD to realize the vision of “data science for everyone”

# House Keeping

- The GitHub repo for this tutorial:
  - <https://github.com/datasciencedojo/IntroToTextAnalyticsWithR>
- Follow Data Science Dojo on social media:
  - Facebook, Twitter, LinkedIn, Instagram, Google+, etc.
- Subscribe to our YouTube channel for more tutorials.
- Don't forget about our 5-day bootcamp!

# Expectation Setting

- I am assuming the following:
  - You have coding experience with R – not an expert, but can hack.
  - You have some math background:
    - The focus will be on practice, not theory – some math is inevitable.
- This is an *introduction* to text analytics:
  - The subject is very, very broad.
  - The focus will be the creation of predictive classification models:
    - The 20% of tools/techniques that can be used in 80% of projects.
  - We will use decision tree ensembles.