“Sharpening the Saw” [[https://youtu.be/pxKxL6RFgjI](https://youtu.be/pxKxL6RFgjI" \t "_blank)]

* 7 Habits of Highly Effective People
* Rebounding from a “Death March”
* Review of Progress to date:
  + Rgui
  + Rgui + Notepad
  + TXT files hackathon
  + TXT files functions
  + RStudio + sourcing (remote/local) using Version Control
  + RNotebook (html) and RMarkdown (pdf using Latex)
* Apprenticeship Model
  + Instructor Role vs. Trainee Role
    - Instructor: May I Coach You?
    - Trainee: On the Court vs. In the Stands
  + 3-Person Problem (Everyone is Learner/Teacher)
  + Legitimate Peripheral Participation
  + Self-directed, Life-long Learner (Problem Solver)
* Details about Instructor
  + Educational Training
  + Life Experience
  + Hybrid: Geek/Manager (Intrinsicity)
  + Personality Disposition (Teller, Risk Taker, Driver)
* Review of Primary Objective: Proficient Data Analyst
  + How is Different from Data Scientist (Liaison)?
  + Four keys to progression from Novice -> Developing -> Mastery
    - Learn the language of data analytics
    - Practice data analytics: from zero to hero
    - Iterate to improve efficiency
    - Learner becomes Teacher
  + Remember, Understand, Apply, Analyze, Evaluate, Create
  + Learning Outcomes
* Review of Task Development to Date
* Re-engage after Reflection/Synthesis

“Upgrade Complete”

* Importance of Notebooks
  + Thought Journal (Non-digital)
  + RNotebooks (Digital)
* Data Provenance
  + Is every stage documented and reproducible?
  + Data Collection
  + Data Cleansing/Merging
  + Sandbox
  + Findings Related to Questions of Interest
    - Pre-hoc Questions
    - New Questions
  + Evaluate Key Findings
  + Prepare Draft of Report
  + Prepare Auxiliary Materials to Support Report
  + Visualizations to Summarize (the power of one)
  + Put Yourself in Shoes of Audience
  + Prepare Final Report
* Proficiency is a Function of Time
  + Reusable Code, if Documented, Saves Time
  + Stream of Conscious
  + Human Performance Theory (Usain Bolt)
* humanVerseWSU
  + Human-Readable Function Names
  + Manager sees “Function Name” not Minutia in Report
  + Evolution of Code:
    - Hackathon
    - Functions-1
    - Functions-2
    - Functions-3 with Documentation
  + Personality Example
  + IMDB Progress Report
* Goal Alignment
  + Fully-function Data-analytics Work Environment
  + STARS you can put on your resume:
    - Situation, Task, Action, Results, Since-Then
    - Start-to-Finish Data Collection -> Report
      * Project 1: Primary Data
      * Project 2: Secondary Data
* Exposure To Various Topics: Suggestions Are Welcome

“SUGGESTED CONTENT”

* Walkthrough of WEEK 00-03 with some Explanation by Instructor
* Histogram w/ Kernel-density estimation w/Normal Curve w/Normality Test
* Optimal Abstraction: Pendulum of Experience & Code Usage

“DISTANCE and CORRELATION”

* Write a Custom Function for computing Manhattan Distance between two points (a,b) where each point has two Cartesian elements (e.g., x,y or latitude,longitude). Extend the function to account for angle of rotation of latitude,longitude inputs (e.g., Chicago city blocks are probably true north-south which means a zero degree offset; Manhattan describes its blocks using uptown/downtown, so what would that angle be?)
* Write a similar function for traditional 2-D Euclidean Distance. Extend the function for a traditional n-D Euclidean Distance. (x,y becomes: x,y,z, …) Think about how the Euclidean formula would be different from an absolute-value formula.
* Write a similar function for a Crow-Flies Calculation (assuming latitude,longitude on a Perfectly Spherical Planet). Answer with 65 votes is a nice starting point: https://bit.ly/2ZxI3ym
* Apply the built-in “dist” function (with default parameters) to the cleansed personality dataset, building a pair-wise distance table. Compare that table to “corr” of the same data. What would change if you standardized the dataset with zscores? Row-wise? Column-wise? Both (in order: row, then column)? Do the distances change? Do the correlations change? [Ideal: left column is distance table, right column is correlation table; each row is how you computed: raw, row-wise, column-wise, both]
* Find other built in distance functions (e.g., Haversine, Great Circle). Benchmark how well your “Crow-Flies” function performs against these other latitude,longitude functions.
* Find some functions that compute the distance between two strings: Jaro-Winkler and Levenshtein are good starting points. Compare the distance between the following words in a pair-wise fashion: TRIANGLE, TRIANGLES , GNARLIEST, RESLATING
  + Maybe take a look at this: http://md5.mshaffer.com/games/wordGame/js/etsgirlan.js
* RELAX with some demo-games I built in 2006-ish:
  + <http://md5.mshaffer.com/games/wordGame/wordGame.html>
  + <http://md5.mshaffer.com/games/sumGame/sumGame.html>
  + <http://md5.mshaffer.com/games/concentration/concentration.html>