**5.3 Data Science Process**

**Close Knowledge Gaps in the Data Science Process**

Recall that the data science process is as follows:

1. Import – Get the data from an original source
2. Store – Store the data into a data store
3. Extract – Get the data from the data store
4. Organize – Organize the data into a usable subset
5. Tidy – Scrub / Clean the data
6. Transform – Change data into something your program will understand
7. Visualize – Descriptive statistics exploration
8. Model – Statistics + Machine Learning + Experiment Design
   1. Where steps 8 & 9 encompass: A. Feature engineering – Variables to use
   2. B. Metric selection – How is model evaluated
   3. C. Algorithm selection – What algorithms to use
   4. D. Parameter optimization – Tune model/algorithm
   5. E. Deployment – Get it into production
   6. F. Evaluation – Did it work? If so, how well?
9. Coding – Machine Learning + Algorithms + Code
   1. Where steps 8 & 9 encompass: A. Feature engineering – Variables to use
   2. B. Metric selection – How is model evaluated
   3. C. Algorithm selection – What algorithms to use
   4. D. Parameter optimization – Tune model/algorithm
   5. E. Deployment – Get it into production
   6. F. Evaluation – Did it work? If so, how well?
10. Understand – Understand insights achieved
11. Communicate – What was the approach, what worked, what didn’t work, what assumptions were made, what would you do different, why do insights matter
12. Next Steps – Where you go from here
13. Document – Code + Thought process + Replication steps

What you’ll want to do now is figure out what small projects you can do with the key words from the pairwise-comparison-winning jobs and the specific steps in the data science process. Remember that you are setting up the structure for your data science portfolio home page based on the data science process so it should be evident to you what gaps in your list you have to fill.

You also want to choose projects that will strengthen your weak skills.

Later when you’ve filed out all the gaps, you’ll want to start again by adding even stronger projects to specific sections where you feel there still exists some weakness. As you develop your skills you’ll find you have a tendency to do very well in some areas. You should also do projects in these areas to make your strengths even stronger.