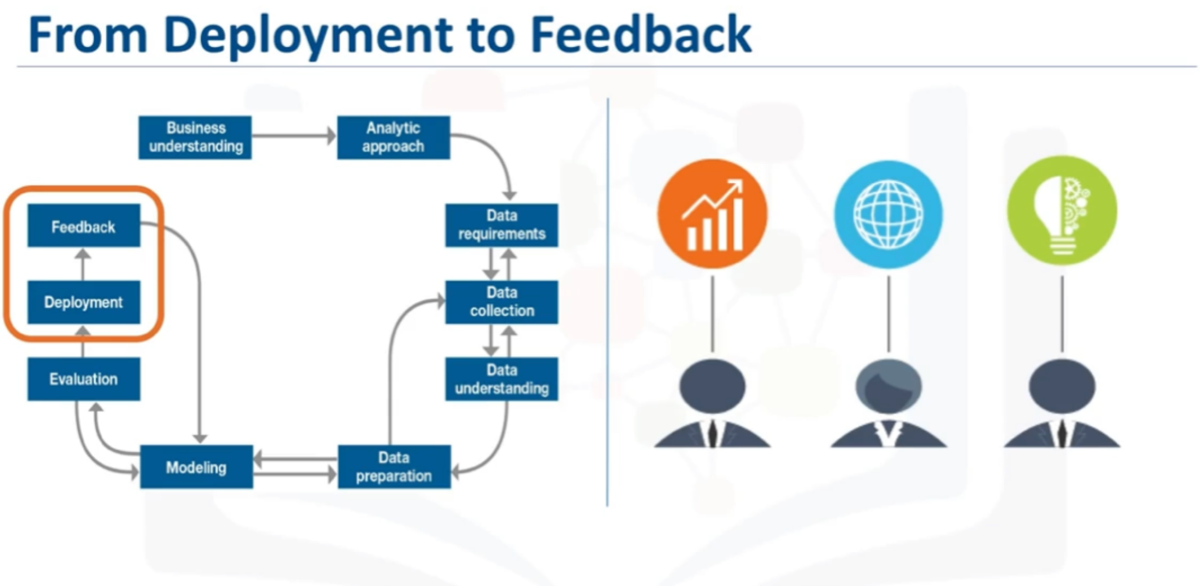
Week 3

#Data Science/3 - Data Science Methodology#

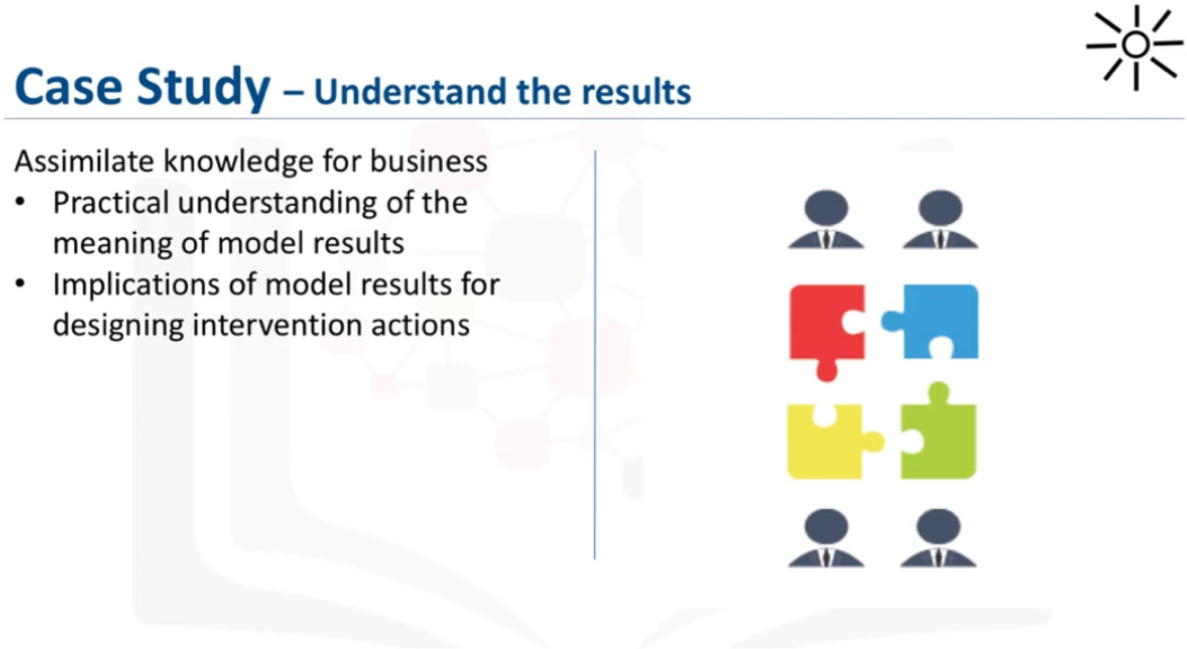
From Deployment to Feedback

Deployment

While a data science model will provide an answer, t he key to making the answer relevant and useful to address the initial question, involves getting the stakeholders familiar with the tool produced. In a business scenario, stakeholders have different specialties that will help make this happen, such as the solution owner, marketing, application developers, and IT administration.

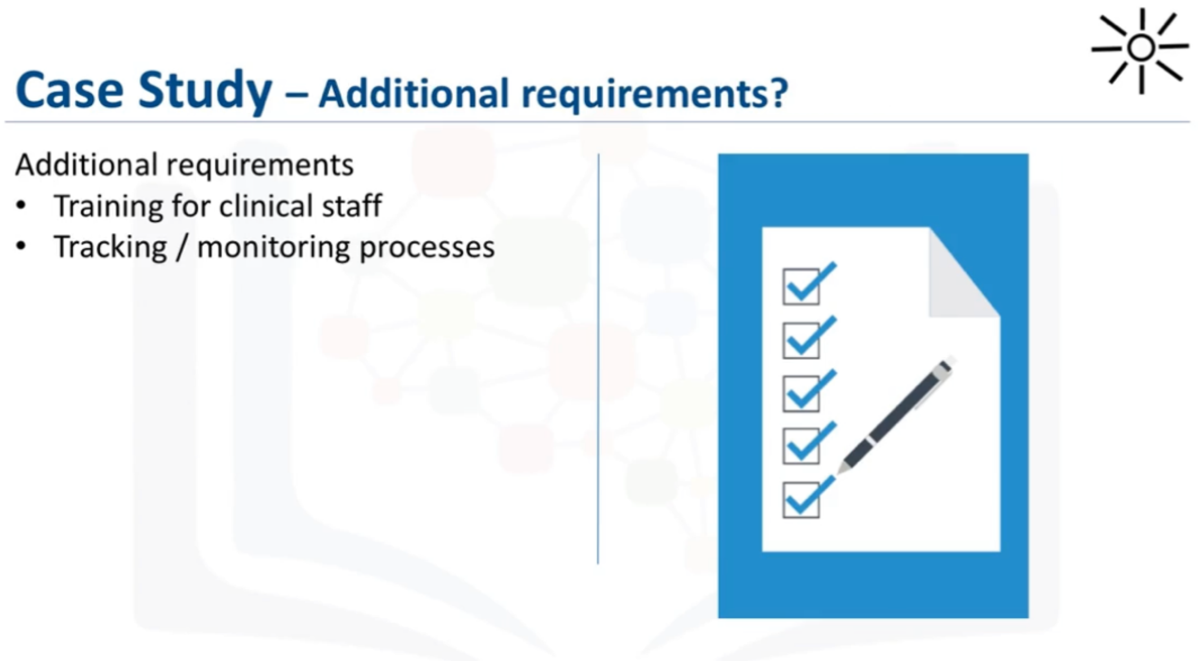


Once the model is evaluated and the data scientist is confident it will work, it is deployed and put to the ultimate test. Depending on the purpose of the model, it may be rolled out to a limited group of users or in a test environment, to build up confidence in applying the outcome for use across the board.
So now, let's look at the case study related to applying Deployment.

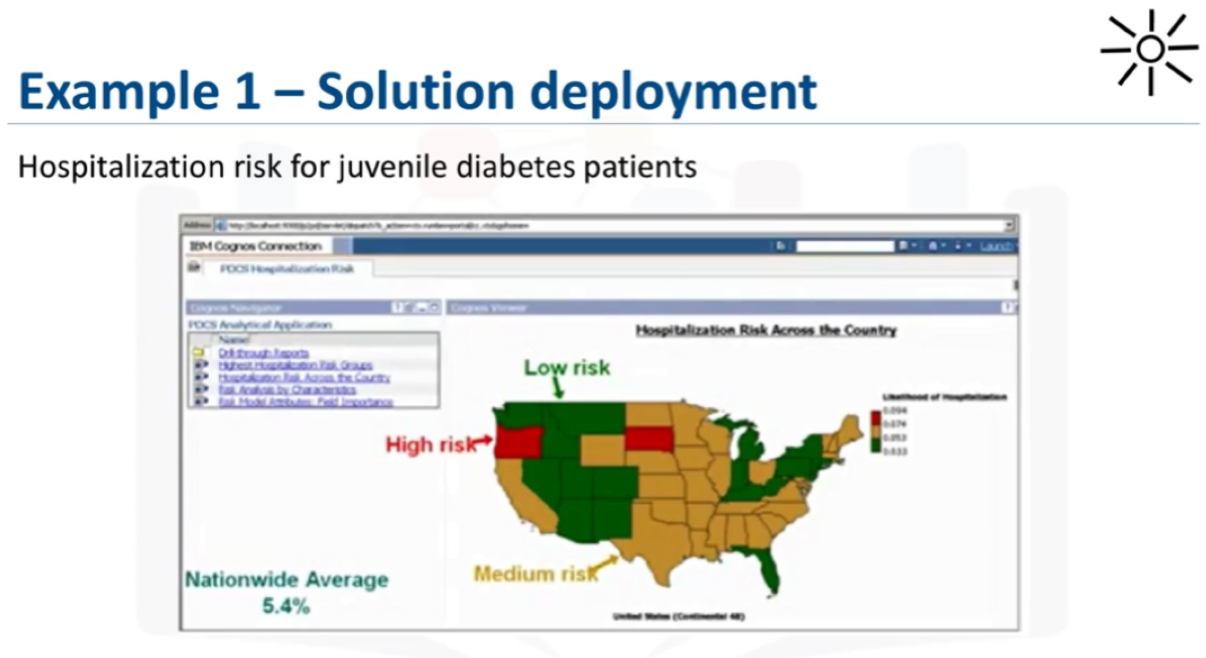


In preparation for solution deployment, the next step was to assimilate the knowledge for the business group who would be designing and managing the intervention program to reduce readmission risk. In this scenario, the business people translated the model results so that the clinical staff could understand how to identify high-risk patients and design suitable intervention actions. The goal, of course, was to reduce the likelihood that these patients would be readmitted within 30 days after discharge.

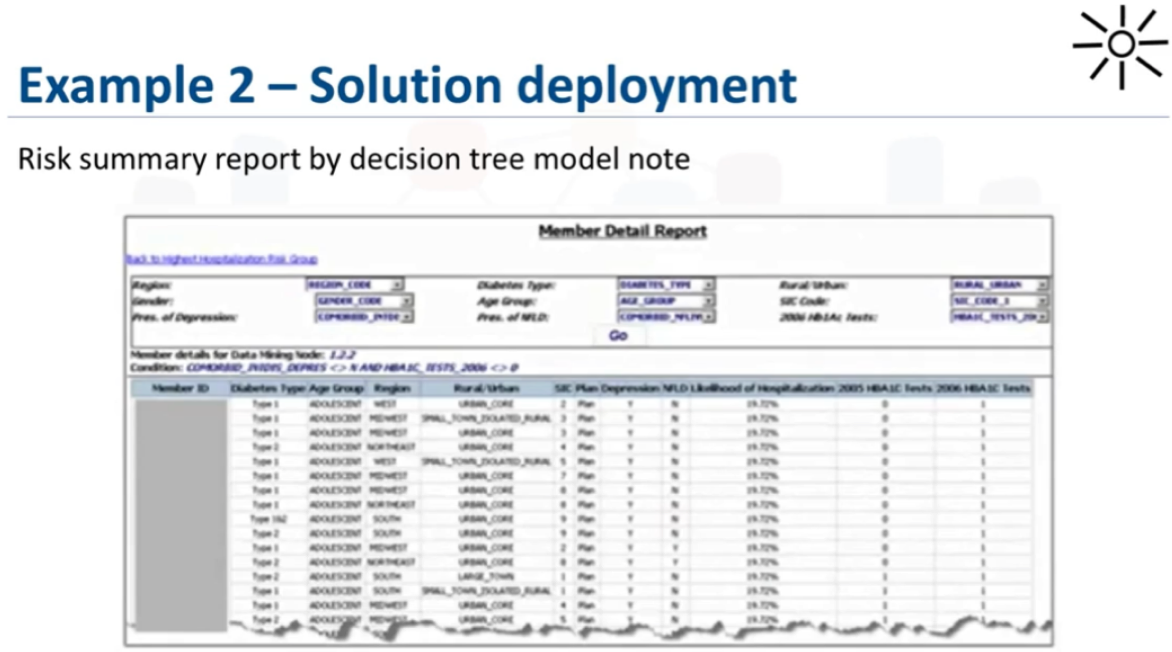
During the business requirements stage, the Intervention Program Director and her team had wanted an application that would provide automated, near real-time risk assessments of congestive heart failure. It also had to be easy for clinical staff to use, and preferably through browser-based application on a tablet, that each staff member could carry around. This patient data was generated throughout the hospital stay. It would be automatically prepared in a format needed by the model and each patient would be scored near the time of discharge. Clinicians would then have the most up-to-date risk assessment for each patient, helping them to select which patients to target for intervention after discharge.



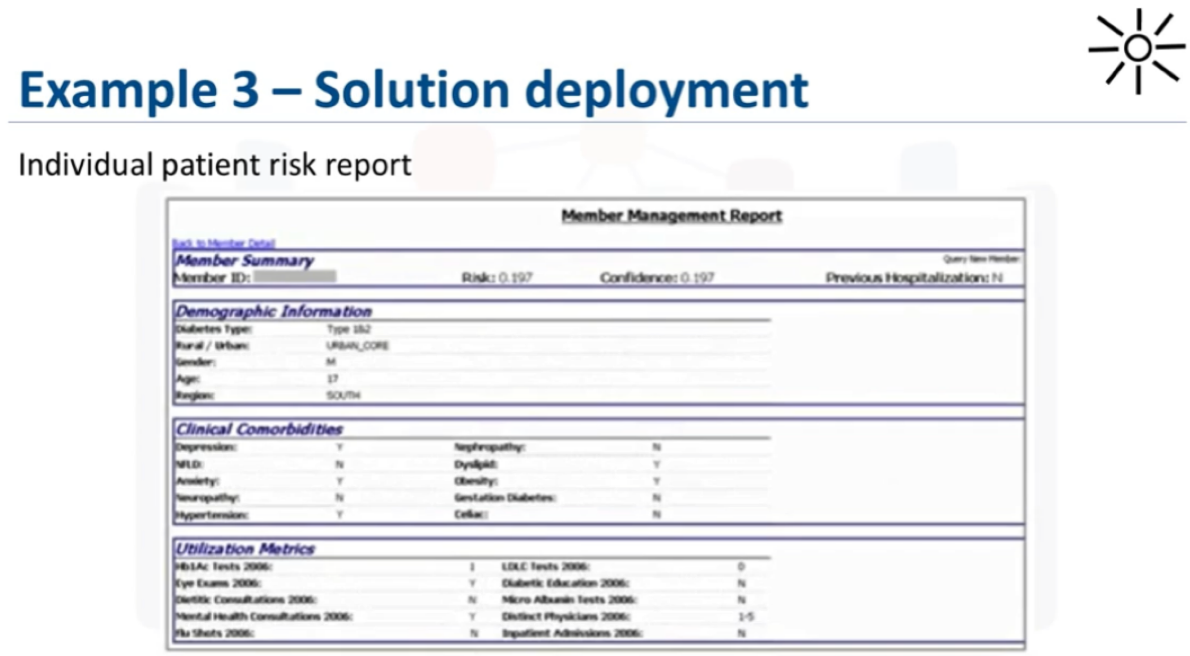
As part of solution deployment, the Intervention team would develop and deliver training for the clinical staff. Also, processes for tracking and monitoring patients receiving the intervention would have to be developed in collaboration with IT developers and database administrators, so that the results could go through the feedback stage and the model could be refined over time.



This map is an example of a solution deployed through a Cognos application. In this case, the case study was hospitalization risk for patients with juvenile diabetes. Like the congestive heart failure use case, this one used decision tree classification to create a risk model that would serve as the foundation for this application. The map gives an overview of hospitalization risk nationwide, with an interactive analysis of predicted risk by a variety of patient conditions and other characteristics.



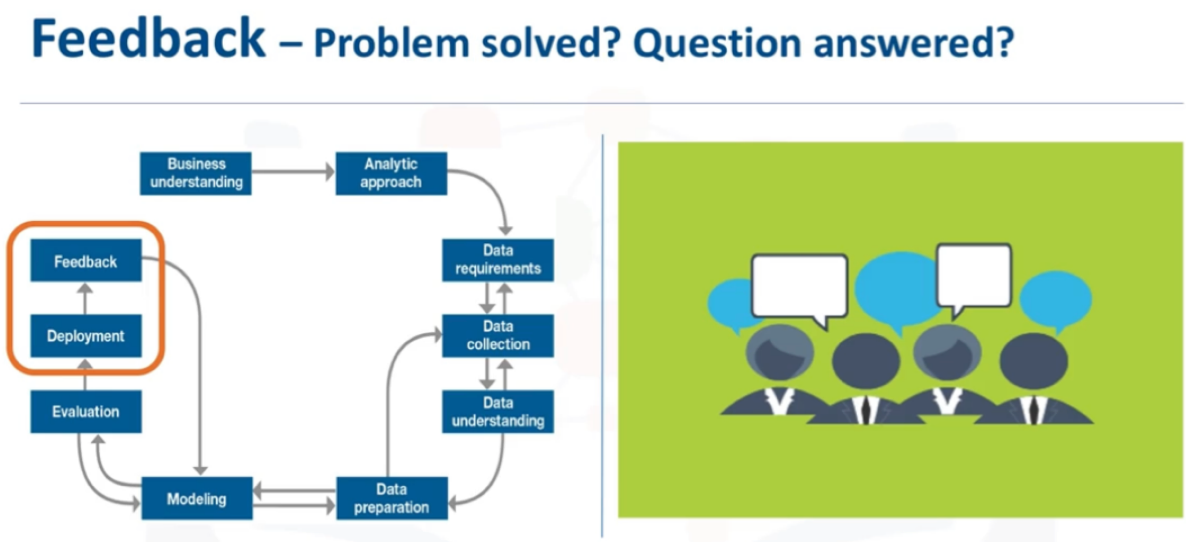
This slide shows an interactive summary report of risk by patient population within a given node of the model, so that clinicians could understand the combination of conditions for this subgroup of patients.



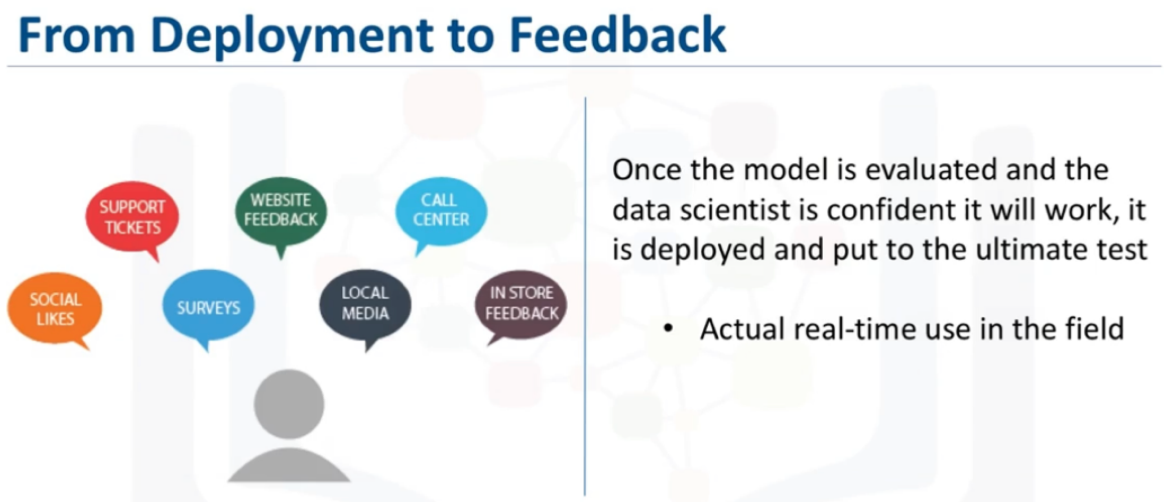
And this report gives a detailed summary on an individual patient, including the patient's predicted risk and details about the clinical history, giving a concise summary for the doctor.

Feedback

Once in play, **feedback from the users will help to refine the model and assess it for performance and impact**. The value of the model will be dependent on successfully incorporating feedback and making adjustments for as long as the solution is required.



Throughout the Data Science Methodology, each step sets the stage for the next. Making the methodology cyclical, ensures refinement at each stage in the game. The feedback process is rooted in the notion that, the more you know, the more that you'll want to know. That's the way John Rollins sees it and hopefully you do too.

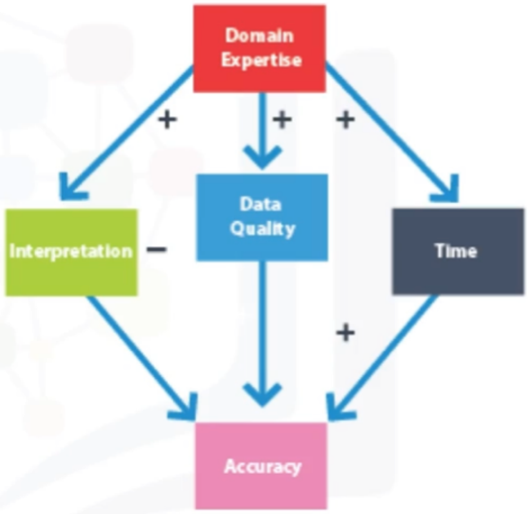


Once the model is evaluated and the data scientist is confident it'll work, it is deployed and put to the ultimate test: actual, real-time use in the field.
So now, let's look at our case study again, to see how the Feedback portion of the methodology is applied.

Case Study – Assessing model performance

Define review process

- To measure results of applying the risk model to the CHF patient population
- Track patients who received intervention
  - Actual readmission outcomes
- Measure effectiveness of intervention
  - Compare readmission rates before & after model implementation



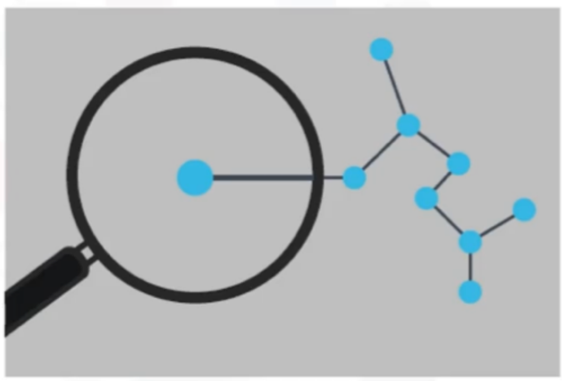
The plan for the feedback stage included these steps:

- First, the review process would be defined and put into place, with overall responsibility for measuring the results of a "flying to risk" model of the congestive heart failure risk population. Clinical management executives would have overall responsibility for the review process.
- Second, congestive heart failure patients receiving intervention would be tracked and their re-admission outcomes recorded.
- Third, the intervention would then be measured to determine how effective it was in reducing re-admissions. For ethical reasons, congestive heart failure patients would not be split into controlled and treatment groups. Instead, readmission rates would be compared before and after the implementation of the model to measure its impact.

Case Study – Refinement

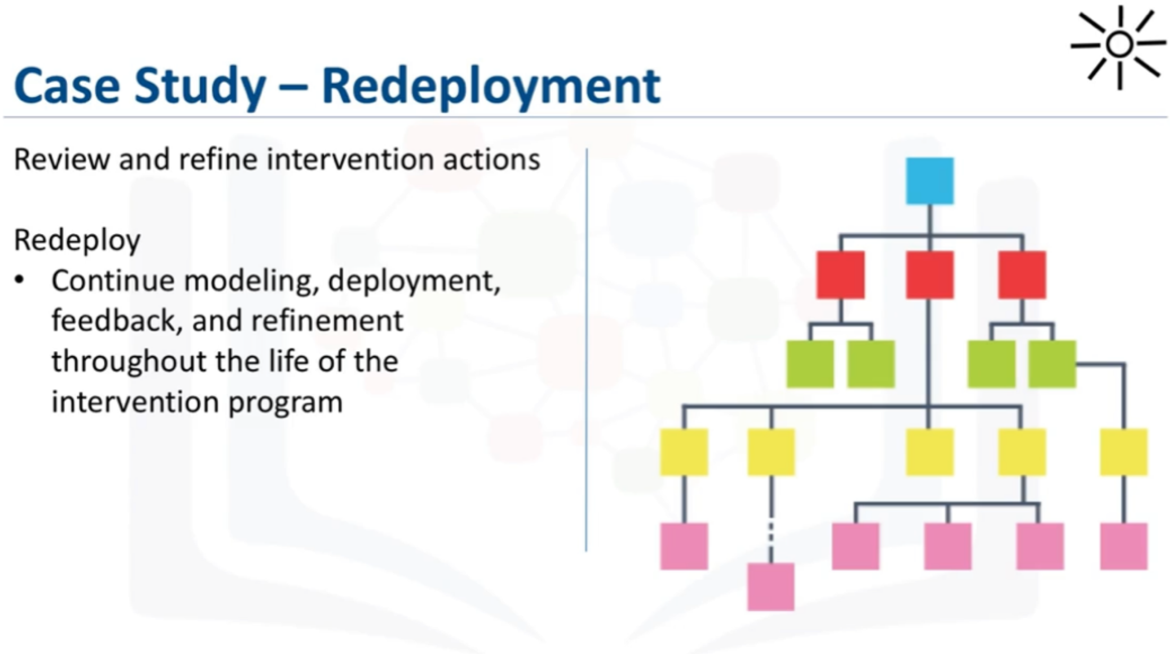
Refine model

- Initial review after the first year of implementation
- Based on feedback data and knowledge gained
- Participation in intervention program
- Possibly incorporate detailed pharmaceutical data originally deferred
- Other possible refinements as yet unknown



After the deployment and feedback stages, the impact of the intervention program on re-

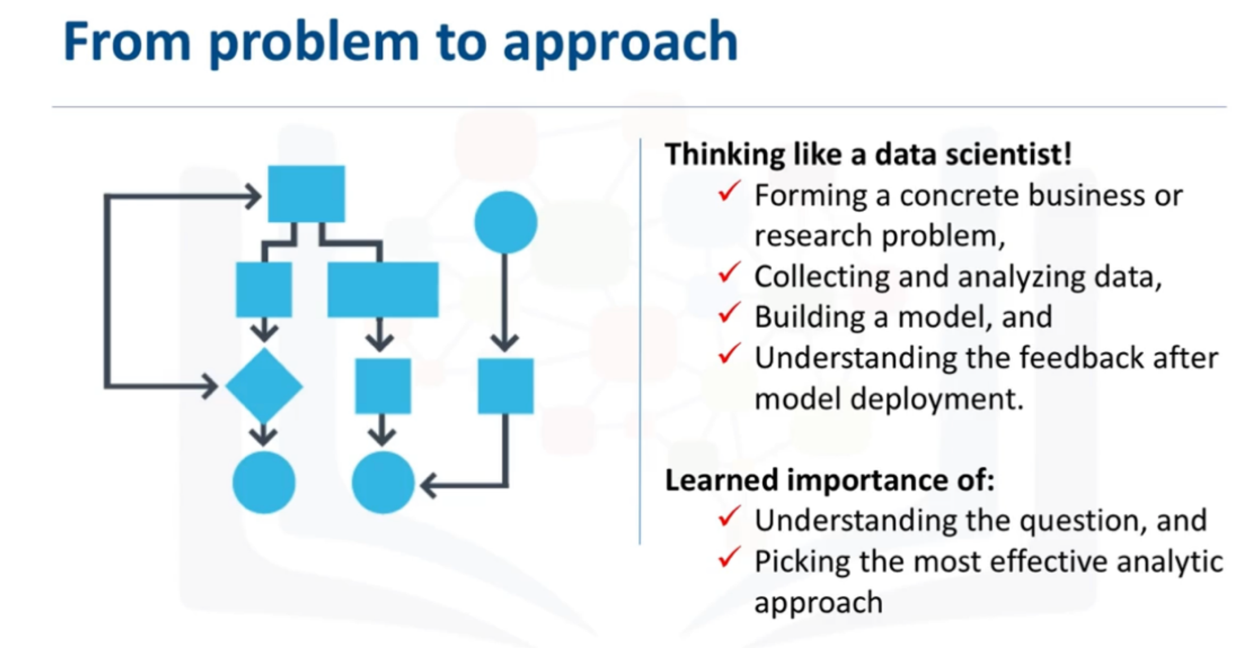
admission rates would be reviewed after the first year of its implementation. Then the model would be refined, based on all of the data compiled after model implementation and the knowledge gained throughout these stages. Other refinements included: Incorporating information about participation in the intervention program, and possibly refining the model to incorporate detailed pharmaceutical data.
If you recall, data collection was initially deferred because the pharmaceutical data was not readily available at the time. But after feedback and practical experience with the model, it might be determined that adding that data could be worth the investment of effort and time.



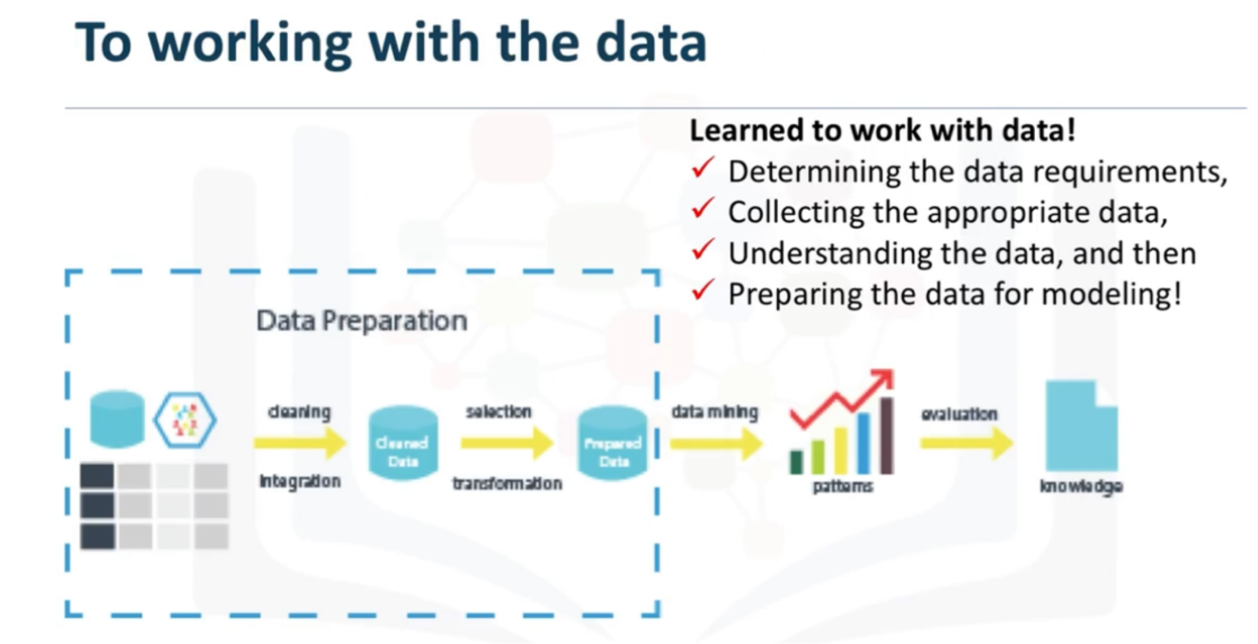
We also have to allow for the possibility that other refinements might present themselves during the feedback stage. Also, the intervention actions and processes would be reviewed and very likely refined as well, based on the experience and knowledge gained through initial deployment and feedback. Finally, the refined model and intervention actions would be redeployed, with the feedback process continued throughout the life of the Intervention program.

Course Summary

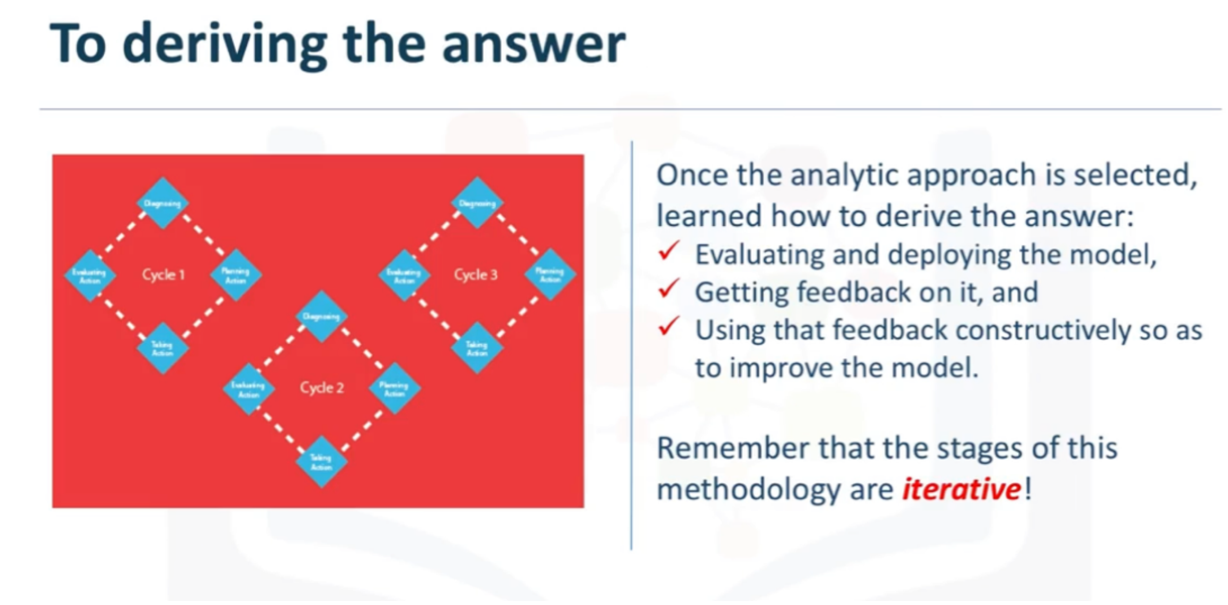
You've learned how to think like a data scientist, including taking the steps involved in tackling a data science problem and applying them to interesting, real-world examples.



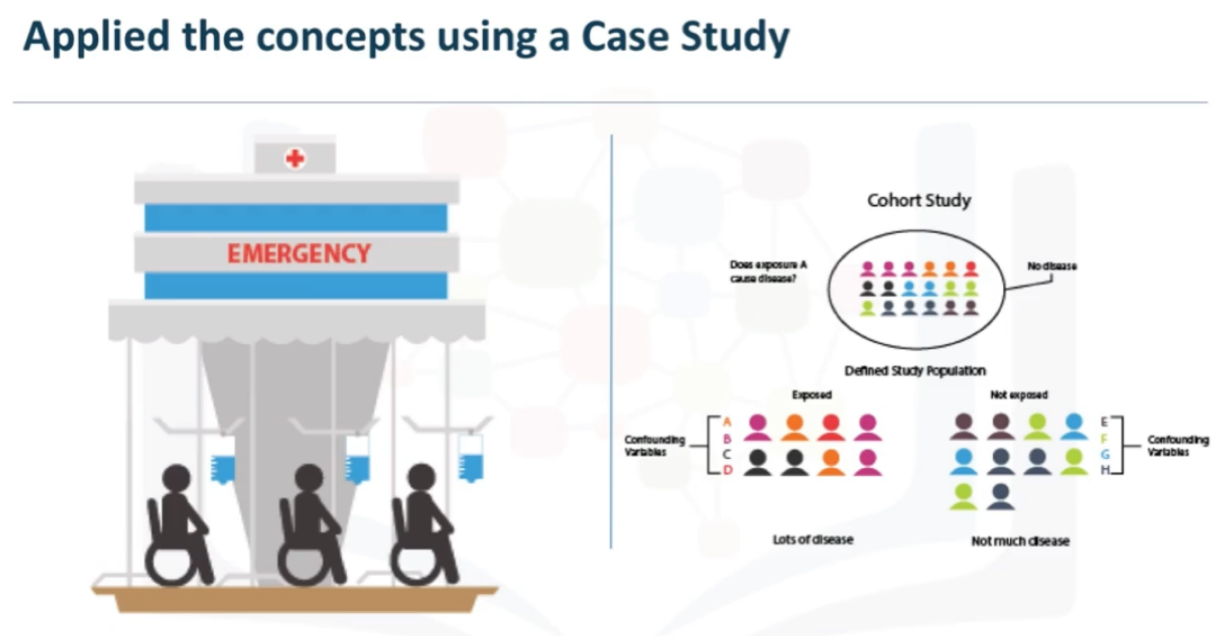
These steps have included: forming a concrete business or research problem, collecting and analyzing data, building a model, and understanding the feedback after model deployment. In this course, you've also learned methodical ways of moving from problem to approach, including the importance of understanding the question, the business goals and objectives, and picking the most effective analytic approach to answer the question and solve the problem.



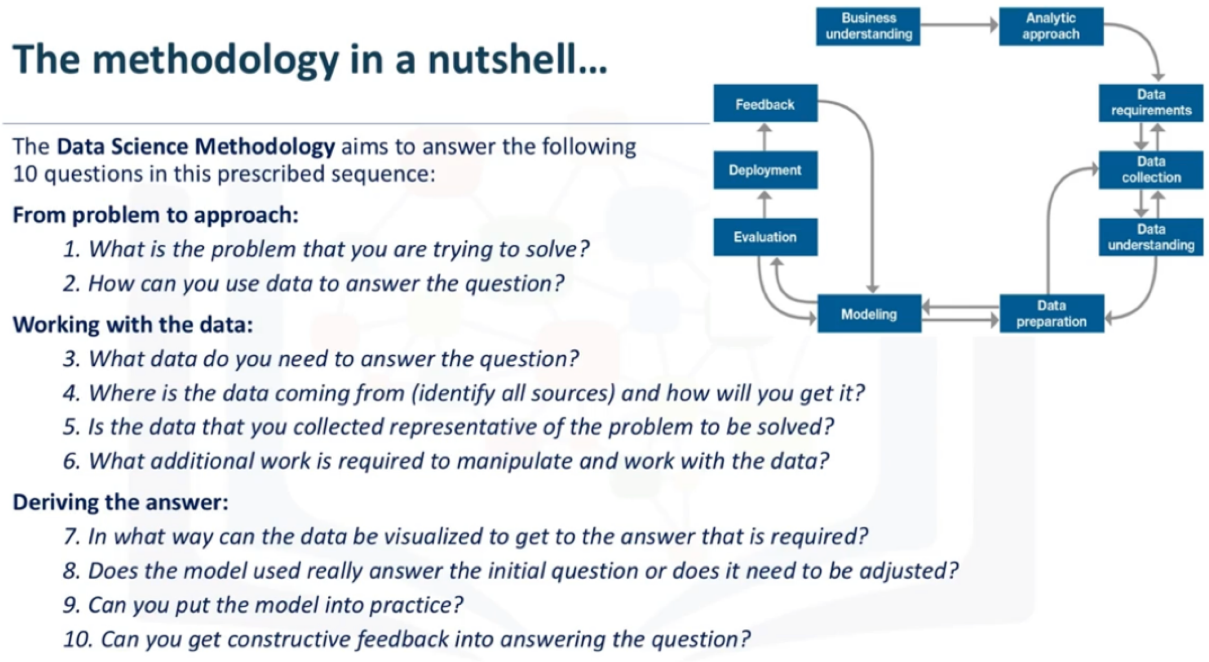
You've also learned methodical ways of working with the data, specifically, determining the data requirements, collecting the appropriate data, understanding the data, and then preparing the data for modeling!



You've also learned how to model the data by using the appropriate analytic approach, based on the data requirements and the problem that you were trying to solve Once the approach was selected, you learned the steps involved in evaluating and deploying the model, getting feedback on it, and using that feedback constructively so as to improve the model. Remember that the **stages of this methodology are iterative**! This means that the model can always be improved for as long as the solution is needed, regardless of whether the improvements come from constructive feedback, or from examining newly available data sources.



Using a real case study, you learned how data science methodology can be applied in context, toward successfully achieving the goals that were set out in the business requirements stage. You also saw how the methodology contributed additional value to business units by incorporating data science practices into their daily analysis and reporting functions. The success of this new pilot program that was reviewed in the case study was evident by the fact that physicians were able to deliver better patient care by using new tools to incorporate timely data-driven information into patient care decisions.



And finally, you learned, in a nutshell, the true meaning of a methodology! That its purpose is to explain how to look at a problem, work with data in support of solving the problem, and come up with an answer that addresses the root problem. By answering 10 simple questions methodically, we've taught you that a methodology can help you solve not only your data science problems, but also any other problem.

Your success within the data science field depends on your ability to apply the right tools, at the right time, in the right order, to the address the right problem. And that is the way John Rollins sees it!