**Deliverable 1 – Health Analytics Project Proposal**

Due: February 2, 2020 at 11:59 PM UTC-12

The Individual Health Analytics Project is an opportunity to conceive, design, and execute a health data research project (Estimation Study) using the skills you will learn in the course.   
([Analytics FAQ](https://gatech.instructure.com/courses/99164/pages/analytics-faq) | [Student FAQ](https://gatech.instructure.com/courses/99164/pages/student-faqs))

**Getting Started**

You will be asking a clinical question that can answered with the types of data available in healthcare related dataset.  We will provide a claims dataset (SynPUF) for your use if you are interested, however you are free to find external datasets for your project. The examples we provide for this project will be based on this dataset.

These data types available in our claims dataset are:

* Demographics
* Visits
* Conditions
* Medications
* Procedures
* Observations
* Death

To find out specifics about what data are contained in SynPUF, look at the [Atlas Data Sources](http://gt-health-analytics.us-east-1.elasticbeanstalk.com/#/datasources/CMSDESynPUF23m/drugera) reports.

**Research Proposal**

The first part of the project is developing your research proposal.  The proposal is not technical but rather a descriptive summary of your planned project.  The proposal comprises the following components:

* **Clinical question**
  + What is the question you are trying to answer?
  + Why is this question significant or interesting?
* **Background**
  + Gather information on your topic from the medical literature (eg Google Scholar)
  + Be sure to gather details about each aspect of your question.
  + Include references
* **Describe your cohorts** (not using code, but in words)
  + Target cohort
  + Comparator cohort
  + Outcome cohort
* **Additional comments**
  + Optional, a place for any comments related to analysis decisions. Not required but highly encouraged. Feel free to share code here etc.

**How to come up with a question?**

There are many sources of questions.  You can look in the news, in medical journals, talk to your family, get ideas from your TAs, or just think about health experiences you have had.  Do you ever wonder if a medication really works?  If you are more likely to get side effects from one drug than another?  Does physical therapy help ?  Do old people get addicted  to opioids as much as young people?  Do diabetes medications lower risk of heart attack?  So many possible questions!

Note, your study should be a comparison study, but it's important to keep in mind that the absence of something is often a comparator.  In other words, a clear comparison would be drug X vs drug Y.  But even if you asked the question "Do diabetics have an increased risk for vision loss?", which may not seem like a comparison study at first, you are actually asking "Do diabetics have an increased risk for vision loss compared with patients without diabetes?"  So really almost any kind of question can be framed as a comparison study.

**An Example (okay, I admit I** [**published a paper**](https://onlinelibrary.wiley.com/doi/full/10.1111/epi.13828) **on this)**

* **Clinical Question**
  + In patients with seizures, does Keppra increase the risk for angioedema compared with Dilantin?
* **Background**
  + In late 2015, the U.S. Food and Drug Administration (FDA) noted a potential safety signal in the Federal Adverse Event Reporting System (FAERS) associating the anticonvulsant levetiracetam (Keppra) with increased risk for angioedema. Anticonvulsant medications are known to be associated with a wide variety of cutaneous reactions, ranging from mild rash to severe hypersensitivity syndromes. Angioedema is a rare but potentially fatal adverse reaction manifested by a rapid swelling of the face, mouth, tongue, and throat.
  + The anticonvulsant phenytoin (Dilantin) has been on the market for over 50 years , is commonly prescribed, and has no documented risk of angioedema.
  + The study's aim is to determine if Keppra increases the risk of angioedema compared with Dilantin.
* **Cohorts**
  + Target cohort
    - Patients with a history of seizures who took Keppra and have at least 60 days of follow up
  + Comparator Cohort
    - Patients with a history of seizures who took Dilantin and have at least 60 days of follow up
  + Outcome Cohort
    - Patient who have a diagnosis of angioedema
* **Additional Comments**
  + Any patient who had angioedema prior to getting Keppra or Dilantin should not be included in this study because it is unlikely that these drugs were the cause of the angioedema
  + We will follow the patients for 365 days after starting their medication to look for the outcome

This example is obviously a specialized medical question.  Yours do not need to deal with such rare things!  But it is important to note that this uses very common types of data (conditions, drugs, etc) that are readily available in claims data.

**Submission**

Submit your proposal in PDF form.  It should be no more than 3 pages (excluding images and references).