





# **Insight Health Data Science Fellows Program**

is an intensive, seven week postdoctoral training fellowship that bridges the gap between academic research and a career in health data science. Based in Boston and Silicon Valley, the program enables scientists to learn the industry-specific skills needed to work in the growing field of big data at leading healthcare organizations.

# Your Bridge to a Career in Health Data Science

Are you a PhD candidate, postdoctoral researcher, or medical doctor looking to transition into a career in industry? Do you want a career that truly leverages your quantitative experience in a fast-growing, in-demand field that is making a positive impact in the world?

Top companies in many fields are hiring health data scientists to help them glean insights from the terabytes of data that they collect everyday. While the amount of data produced and stored is growing exponentially, there is a severe shortage of talent to analyze this data and extract valuable insights from it.

The Insight Health Data Science Fellows Program is a postdoctoral training fellowship that bridges the gap between academia and health data science. Insight is a seven week, intensive program that teaches researchers how to apply their existing analytical skills to big data. The program consists of a project-based, industry-driven learning approach designed to train academics in many of the cutting edge data science tools and healthcare practices used in industry. Immediately following the program, Fellows interview at leading organizations in Boston and the surrounding area, where they are likely to receive multiple job offers.

#### Insight in a nutshell:

- 1. 7 week, full-time, postdoctoral health data science training fellowship in Boston and Silicon Valley.
- 2. Need-based scholarships available to help cover living expenses.
- 3. Self-directed, project-based learning (*no classes!*) under the guidance of top industry health data scientists.
- 4. A group of smart people who are excited about working on interesting problems while having a positive impact.
- 5. Interview at top health companies and organizations immediately following the program.

**INSIGHT** HEALTH DATA SOIENCE



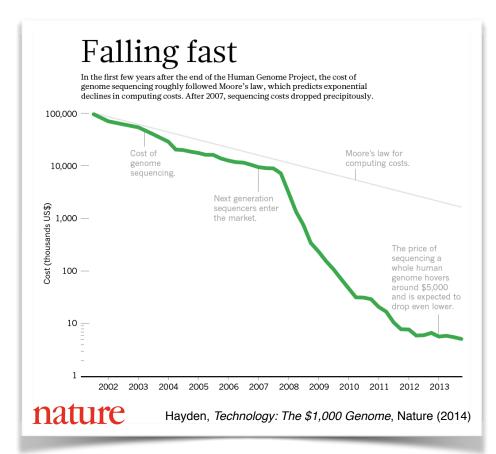




## What is Health Data Science?

Across the health and life sciences sector, servers are overflowing with data: genome sequences, drug screens, clinical trial data, electronic medical records, claims data, longitudinal studies, and even quantified-self data. This information is a form of "big data" not only for the sheer size but also its complexity, diversity, and time dependence. Health data scientists analyze this data to find more effective treatments for a given disease, increase quality of life using personalized lifestyle recommendations, discover new types of cancer treatments, provide realtime health information to doctors or families, and link environmental factors to disease and corresponding treatment options.

"The McKinsey Global Institute estimates that applying big data strategies to better inform decision making could generate up to \$100 billion in value annually across the US health-care system, by optimizing innovation, improving the efficiency of research and clinical trials, and building new tools for physicians, consumers, insurers, and regulators to meet the promise of more individualized approaches<sup>1</sup>."



<sup>1</sup> McKinsey, How big data can revolutionize pharmaceutical R&D, (2013)

How is data science being applied to health? Here are just a few examples:

**Clinical Trials:** Patients are being automatically identified for clinical trials based on multiple sources of data including electronic health records, claims data, and genetic information. Selectively targeting specific populations will lead to smaller, faster, cheaper, and more effective clinical trials.

**Healthcare Analytics:** Predictive modeling of electronic health records and claims data is identifying hot spots within the healthcare system. This approach was recently used to successfully identify the tiny fraction of individuals that accounted for 30% of the total costs for one hospital system<sup>2</sup>.

**Genomics:** The cost of sequencing the whole and fraction of human genome has dropped dramatically in recent years, which is making the technology more accessible and common in diagnosing diseases. DNA sequencing is being used in precision medicine that pinpoints specific cancer mutations and recommend optimal treatment based on the patient's genomic profile.

**Quantified Self:** Petabyte of quantified self data are generated and streamed into the cloud from 1/10 Americans who wear fitness devices every day. Self-reported health status are being collected on over 2,500 mobile apps. Data scientists at various organizations just began to use this rich data to monitor disease progression, diagnose early onset of health conditions, and recommend healthy lifestyles.

**Healthcare Delivery:** Medication non-adherence is a growing concern to clinicians, healthcare systems, and other stakeholders due to its high prevalent and associated with adverse outcomes and higher costs of care. Internet-of-Things technologies monitor patient behaviors and the data they collect are being used to generate personalized intervention to increase patient engagement.

**Drug Development:** The biological connection between proteins, genes, pathways, drugs, and diseases is revealing new information about diseases or treatments. Mining existing biological knowledge and creating novel discovery models dramatically reduces time and cost of drug development process.

In order to derive meaningful value from health data, biotech companies, pharmaceutical companies, startups, research institutes, payors, medical record providers, and hospitals are hiring data scientists focusing on health. These health data scientists will be at the forefront of developing the next generation of analytical tools, performing complex analysis, and creating the data infrastructure that will ultimately improve the lives of patients.

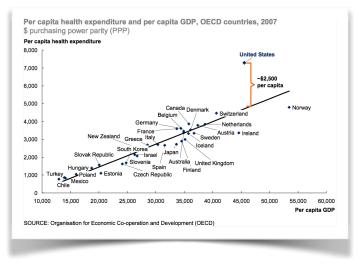
<sup>&</sup>lt;sup>2</sup> Gawande, The Hot Spotters, The New Yorker, (2011)

# How Big is the Opportunity for Health Data Scientists?

"The number of challenges that one sees at the intersection of the life sciences and technology is quite vast. I think the number of opportunities is even greater<sup>3</sup>."

– Anthony Philippakkis, Chief Data Officer, The Broad Institute

Many life sciences and healthcare companies have acknowledged that they are drowning in data<sup>4</sup>. These companies realize that health data science is no longer optional; it is necessary to remain competitive and improve patient care.



The U.S. spends more per person on healthcare than any other nation (see graph on left)<sup>5</sup>. As healthcare expenditures grow at an unsustainable rate, health data science is emerging as an opportunity to help millions of people and save billions of dollars.

To tackle these challenges, life sciences and healthcare companies are creating health data science roles. Hospitals like Memorial Sloan-Kettering Cancer Center are hiring health data scientists to mine 1.4 million patient records over 25-years to control and cure cancer.

Research institutes like the Broad Institute are hiring health data scientists to develop the next generation infrastructure to integrate petabytes of genomic, proteomic, chemical, and health datasets to enable new discoveries. Pharmaceutical companies like Novartis are hiring health data scientists to find novel personalized approaches to cancer treatment, while Biogen Idec is leveraging their data to predict the impact of novel drugs on patient outcomes. Wearables like Fitbit and Jawbone are hiring health data scientists to investigate behavioral changes and health benefits. In addition, with the help of health data scientists, numerous health startups are starting tackling various challenges in healthcare: optimizing hospital operations, increasing access to care, improving patient-physician communication and even developing novel therapeutics for chronic diseases such as cancer.

<sup>&</sup>lt;sup>3</sup> Mass Technology Leadership Council, Big Data & Life Sciences, (2013)

<sup>&</sup>lt;sup>4</sup> Fortune, *3 Ways Big Data Is Changing Life Sciences*, (2014)

<sup>&</sup>lt;sup>5</sup> McKinsey Global Institute, *Big data: The Next frontier for innovation, competition, and productivity,* (2011)

#### Who Are the Best Health Data Scientists?

"Bring your expertise to healthcare and you'll know you're going to make a difference, at the patient level and at the societal level."

> — George Telthorst, Director, Center for the Business of Life Sciences, Indiana University

The best health data scientists have a passion for solving healthcare challenges with quantitative methodologies. In particular, their backgrounds are in diverse fields such as bioinformatics, public health, medicine, psychology, physics, astrophysics, neuroscience, math, statistics, computer science, engineering, and other data-heavy disciplines.

This diversity is not surprising as data scientists at technology companies, where data science originated, come from similar backgrounds. According to DJ Patil, the Chief Data Scientist for the White house who co-coined the term *data scientist* and served as an Insight mentor and advisor, using "medical and genomic data to build a model of patient-powered research to accelerate discovery and help clinicians select which treatments will work best for which patients" is a major priority of the current administration<sup>6</sup>.

While scientists make great health data scientists, those currently in the industry often take a long and winding road to get there, learning the tools used in big data informally through trial and error. The emerging health data science field requires knowledge about relevant programming languages and databases, statistics and machine learning, life science fundamentals, and healthcare industry practices. This is where the Insight Health Data Science Fellows Program comes in, and why top life sciences and healthcare organizations are helping Insight develop the next generation of leading health data scientists.

<sup>&</sup>lt;sup>6</sup> DJ Patil at Strata + Hadoop 2015

## **Insight Health Data Science Fellows Program**

"Insight gives people a practical path to move into informatics."

Stephen Cleaver, Executive Director of Informatics
Systems, Novartis

As a scientist, you possess many of the fundamental skills necessary to be a great health data scientist: a passion to help others, big picture problem solving, strong quantitative abilities, and experience with statistical analysis. While you have 90% of the foundational skills needed, you are missing the final 10% - experience with healthcare data and the tools and techniques that would enable you to be productive on the first day of your new job as a health data scientist. Companies are growing rapidly and simply cannot afford to "take a chance" on someone who still needs several weeks of on-the-job training just to become a productive team member. As a result, a skills gap exists between the world of academic research and health data science.

That's where Insight comes in. We accept top PhDs, MDs, and postdocs and provide them with the time, space, and resources necessary to get up to speed on the tools, techniques, and models they will need to learn to get hired as a health data scientist and hit the ground running in their new career.

Here's what you need to know about the Insight Health Data Science Fellows Program in a nutshell:

- 1. 7 week, full-time, postdoctoral health data science training fellowship in Boston and Silicon Valley.
- 2. Need-based travel and living expense scholarships available.
- 3. Self-directed, project-based learning (*no classes!*) under the guidance of top industry health data scientists.
- 4. Alumni network of over 500 Insight Fellows who are data scientists and data engineers.
- 5. Interview and get hired at mentor organizations immediately following completion of the program.

## **About Insight**

Insight was founded in 2012 with the launch of the Insight Data Science Fellows Program, a 7-week project-based postdoctoral training fellowship, designed to help science PhDs become data scientists. The goal for Insight programs is to accept top scientists and provide them with the resources they need to become data scientists at innovative companies. Thanks to the hard work of the Fellows and the help of our industry mentors, the program has been a resounding success. There are now over 500 Insight Fellows working as data scientists and data engineers across Silicon Valley, New York City, Boston and other major cities (LA, DC, Seattle etc). We work with every Fellow until they accept a position they are happy with. Fellows who have stuck with the post-Insight job search process have found work in a data-related role within 3-4 months of completing the program, with a majority of Fellows getting one or more offers from Insight mentor companies within 4-8 weeks immediately following the program. They are now working on top data-driven teams including Facebook, LinkedIn, Counsyl, Palantir, Invitae, Jawbone, Memorial Sloan Kettering Cancer Center, Biogen, athenahealth, The Broad Institute as well as at various well-funded startups. You can find out more about the Insight Data Science Fellows here: http://www.insightdatascience.com/ fellows.html



While the Insight Data Science and Date Engineering Fellows Programs are geared towards technology companies, numerous applicants are interested in making a direct impact on human health by working for a life science or healthcare organization, who are looking for highly motivated data scientists with domain knowledge. We launched a dedicated Insight Health Data Science Fellows Program in Boston in July 2015. Since then we've seen great success with all Fellows land on data science and/or health analytics roles within 4 months upon completion of the program. Many mentor companies return to Insight to bring top-notch candidates to their data teams.

Recently, more and more west-coast health companies are requesting to be Insight mentor companies because they need health data scientists to tackle their big-data challenges. Additionally many of our applicants would love to work in Silicon Valley to while making a direct impact in healthcare.

To meet the growing demand from both interested Fellows and organizations hiring health data scientists, we are launching a dedicated Insight Health Data Science Fellows Program in **Silicon Valley** in **September 2016**. The educational model we pioneered with the Health Data Science program in Boston will be used for its sister program: project-based learning under the mentorship of top industry professionals and in-depth conversations with leading experts to learn and understand the challenges facing health data science.

#### Who's Involved?

Insight is an education startup working to bridge the gap between academia and quantitative careers in industry. With seed funding from startup investment funds **Y Combinator**<sup>7</sup>, **SV Angel**<sup>8</sup>, and **Avalon Ventures** and, with participation from leading technology, healthcare, and life sciences companies, we are connecting top analytical talent with some of the most innovative organizations in the world. Mentors for the Insight Health Data Fellows Program, who will be looking to hire for their teams, are health data scientists at:

<sup>7</sup> http://ycombinator.com

<sup>&</sup>lt;sup>8</sup> http://svangel.com, http://economist.com/node/21537967





#### How Does it Work?

The Insight Health Data Fellows Program is a full-time, seven-week postdoctoral fellowship that helps scientists transition from academia to a career in health data science. Office space is included for the duration of the program, and we help international students handle any necessary visa arrangements. Finally, need-based living expense scholarships are available.

During the first week of the program, you are introduced to the field of health data science by participating in round table discussions with leading experts from life sciences and healthcare organizations. In addition, you will gain hands-on experience using the tools, techniques, and best practices for analyzing health data, while brainstorming possible data projects. By the end of the first week, with the input of the mentors and peers, you will select a health-related data problem that you will solve in the subsequent weeks.

Over the next three weeks, you will work on your project, learning the necessary background, technologies, and techniques that you need to solve your health data problem, which will showcase your skills as a health data scientist. In weeks five, six and seven, you will prepare for interviews while presenting your project to the various mentor organizations that you're interested in interviewing with. Formal interviews begin after the end of the program. After completing all the interviews, 4-8 weeks after the end of Insight, you will likely have one or more job offers from leading organizations and be ready to start your career as a health data scientist.

## **Program Details**

The Insight Health Data Science Fellows Program is a seven week, full-time program that consists of the following:

	Program Timelin	Ð	
7 Weeks		4-8 Weeks	
Weeks 1-4	Weeks 5-7		
Data Project	Company Visits	Job Interviews	~

**Intro to Health Data Science** During the first week of the program, top healthcare and life sciences experts will have round table discussions and lead you through the necessary foundational concepts of health data science, giving you a big picture overview of the field and how you can become a great health data scientist. In addition, you will learn about the tools, techniques, and best practices for analyzing multiple sources of health-related data all while brainstorming possible health data science projects. By the end of the first week, with the input from your peers and the mentors, you will select a project that you will work on in the following weeks.

**Health Data Science Project** In weeks 2-4, you'll work exclusively on solving the health data project you identified in week one. The purpose of the project is to showcase your existing data analysis skills by applying them in the context that healthcare organizations deal with on a regular basis. Although you will leverage your existing skill set, it may be your first time using big data technologies or working with different types of healthcare data (e.g., genomic data vs. claims data). The project will serve as the method by which you will learn the technical skills and technologies that are standard in the healthcare industry. These include:

**INSIGHT** HEALTH DATA SOIENCE

1. Life Sciences and Healthcare Software Engineering Best Practices: Learn how to contribute to a large code-base, learn the pros and cons of various open-source big data technologies, and run analyses on a distributed computing system.

2. **Storing and Retrieving Big Data:** Learn how to extract, transform, and load (ETL) data into the appropriate database or distributed data storage system. Learn how to write sophisticated queries to retrieve information needed for analysis.

3. **Integrating and Normalizing Diverse Datasets:** Learn how to combine multiple sources of data in order to tackle difficult questions. Learn how to recognize common issues that arise when integrating and normalizing data.

4. **Statistical Analysis and Machine Learning:** Learn industry best practices for doing basic and advanced statistical analysis on large data sets.

5. **Visualizing and Communicating Results:** Learn how to effectively communicate your findings visually and verbally. Learn how to present datadriven results that a company or organization can act on.

Throughout Insight, there are no grades or other arbitrary proxies used to evaluate your work. Instead, your data analysis project will serve as your "professional portfolio," which you will be able to show to potential employers in order to demonstrate your understanding and proficiency in the subjects tackled.

**Collaborative Learning** While your work will be self-directed, you are never left alone to fend for yourself. Your peers are there to assist you and industry mentors are on hand daily to discuss difficult-to-understand concepts or to help fix bugs. Group discussions and Q&A sessions will also be a regular part of the program, along with an informal collaborative, peer-learning culture being encouraged at all times.

**Mentors** Throughout the program, you will be interacting on a daily basis with Insight mentors, all of whom are leading data scientists, data engineers, health data scientists, or managers from industry. These are the people who work at the very same companies that you will have the opportunity to interview with at the end of the program. This means that you will not only learn about the cutting edge techniques being used at these companies, but you will also get to know the actual practitioners themselves, who are actively working in the field. As a result, you will develop professional relationships with dozens of health data scientists. This is an invaluable professional network that you will be able to draw on throughout your career.

**Interview Preparation** In addition to the project work and collaborative learning, you will get a chance to do practice interviews to prepare for the real thing. These mock interviews will also allow you to practice clearly articulating the experience you developed in your prior work while also developing your health data project at Insight.

**Company Visits & Matching** Starting in weeks 5-7, you will get an opportunity to visit the offices of the companies you're interested in interviewing with and present your project to their health data science teams. Throughout the program you will interact with mentors and participate in round table discussions that will help you decide which companies you are most interested in presenting your health data project to. Most Fellows visit and present at five to ten companies that they are most excited about. The companies then reach out to those Fellows whom they feel would be a good fit for their teams and schedule full interviews.

**Interviews** Starting in week eight, you will begin the interview process with the companies that have reached out to you as a result of your Insight project presentations in the previous weeks. Most Fellows interview with anywhere from 3-6 companies. While the interview process *starts* immediately following the end of week seven of Insight, it usually continues for another several weeks, with most Fellows receiving offers 4-8 weeks after the final day of Insight.

#### **Being an Insight Fellow**

The goal of Insight is to train the next generation of leading health data scientists. To do this, we have created a program that is explicitly designed to walk Fellows directly into fulfilling careers at companies on the leading edge of health data science and analytics. These companies have a very high bar for talent and are only looking for the best possible candidates in any position they hire for. This is why we are setting a very high standard for acceptance into the program and expect entry to be quite competitive.

While top-tier quantitative ability and scientific data analysis experience is necessary, it is not sufficient to be an Insight Fellow. In addition to sheer smarts, we're looking for Fellows who are passionate about improving human health by leveraging big data opportunities. Fellows must also be extremely curious, highly motivated, love learning across a wide range of fields, enjoy collaborating with others, be self-driven, and are excited about the opportunity to make a positive impact in the world.

#### **Responsibilities of Fellows**

As an Insight Fellow, you're given the opportunity to learn from the best health data scientists in Boston for seven weeks. The program is designed to remove as many obstacles as possible that stand between where you are now and becoming a health data scientist. With these benefits, however, come a few responsibilities that you must be comfortable with before choosing to apply for the program:

- Full-time participation is expected in Boston for seven weeks of the Insight program and three to five weeks of interviews immediately following. During the program, you will be required to be at the office from 10am-6pm Monday to Friday. Some days, you will need to stay for mentor company visits ending as late as 8 or 9 p.m., accommodating mentors who can only attend in the evening.
- Completing your Insight project will require independent work above and beyond the structured time. Fellows who have taken full advantage of the program often work late nights and weekends, collaborating with their peers. We are looking for Fellows who are enthusiastic, active participants in this intense program.
- You must intend to take a job as a full-time health data scientist after the program is finished and agree to interview with mentor companies

immediately after the program. If you are interested in working at a company that hasn't participated in Insight, we'd love to help you navigate that process too, but you must let us know in advance of contacting them, so that we can get them involved in the program.

• You agree to become a mentor for future Insight Fellows. In particular, this means coming in for at least three two-hour visits during the subsequent Insight program session, and two additional visits over the following year. Alumni involvement is a cornerstone of the program, and this commitment will help future Fellows learn from your experiences.

The guiding principle of Insight is to be an advocate for the Fellows and create an environment where you can learn and develop into a great health data scientist. All we ask in return is that you give it your all, be fully engaged in the process, and help pass on your learning to the next batch of Fellows, helping to make the Insight community stronger as a result.

#### **Benefits to Fellows**

The Insight program is designed to provide all the training, resources, and connections you'll need to effectively transition from academic research to a career in health data science. Here are some of the benefits of becoming an Insight Fellow:

- Need-based scholarships are also available to help cover living and travel expenses -- our goal is to make sure everyone with the right skills can participate in Insight, regardless of their financial situation.
- Desk space at the Insight offices in Boston or San Francisco during the program.
- Tips and assistance from our staff to help with your living arrangements for the duration of the program.
- Guidance and mentorship from industry professionals at every stage of the program and as you prepare for interviews.
- Mentorship from alumni Insight Fellows whose experience, at Insight and at their current data roles, make them an unparalleled resource to provide guidance and feedback.

- Personalized company matching. We help you figure out which companies are the best fit for you based on our experience and in-depth conversations with the hiring managers. We then help you arrange interviews during the final week of the program.
- Help navigating the negotiation of final employment terms once companies have made their employment offers to you.
- Perhaps most importantly: an unparalleled professional network of data scientists, data engineers, friends and acquaintances. Through the program you will meet and get to know several dozen top health data scientists, who are Insight mentors and alumni, all of whom will be your industry peers. These professional contacts will be an invaluable source of knowledge, advice, career opportunities, and friendship in the years to come.

#### **Frequently Asked Questions**

**Will I get hired after completing Insight?** Yes. We aim for approximately one hiring company participating in each session per Fellow. You will effectively be walked into interviews with a half dozen or more companies all of whom are extremely interested in hiring health data scientists. Most Fellows who completed the job search found work in a data-related role within 3-4 months of completing the program, with a majority of Fellows getting one or more offers from Insight mentor companies in the 4-6 weeks immediately following the program. The entire fellowship is a training program where you are taught to be a productive health data scientist from day one at your new place of work, and to demonstrate this in the interviews. When you combine all of these factors, it is very safe to assume that as long as you put effort in and actively pursue the job search process through to completion, you will receive job offers following the fellowship.

What is the best time to do the Insight Health Data Science Fellows Program? The ideal time to do the program is when you are *less than* four months away from your ideal start date for your full-time job as a health data scientist. In other words, if you will not be able to start your full-time job in industry until more than four months after the last week of Insight, you should consider applying for a later session. Some Fellows have negotiated start dates up to one year away, but these are special exceptions. In general, we recommend taking time off from your PhD or Postdoc when you are 1-3 months away from graduating or finishing your research commitment. This way you can do Insight, interview with the hiring companies, accept a job offer, and then go back for 1-3 months to finish up your research.

**Do you accept Fellows who are not U.S. citizens or green card holders?** Yes. However, you must make sure that you will be able to work in the U.S. from the start date of your full-time employment. For Fellows who are completing (or recently completed) a graduate degree in the U.S., this usually means applying for an Optional Practical Training (OPT)<sup>9</sup> period, after which the hiring company will apply for an H1B visa<sup>10</sup> for you. If you qualify for OPT, feel free to apply for any Insight session, as timing is not an issue. If you do *not* qualify for OPT, then we recommend that you apply for a **January** session of Insight, to ensure that the hiring companies can apply for an H1B visa for you before the April 1st filing date.

<sup>9</sup> http://en.wikipedia.org/wiki/Optional Practical Training

<sup>&</sup>lt;sup>10</sup> http://en.wikipedia.org/wiki/H-1B\_visa

What type of impact can I expect to make? Many of the companies that are involved in Insight and at which you'll have the opportunity to work, are truly changing the world for the better. They're discovering new drugs for cancer. They're diagnosing diseases by sequencing and interpreting genetic information. They're creating new sensors to automatically monitor our health. Many of them are integrating healthcare data in novel ways to track drug response and prevent adverse drug reactions. Still others are working on behavioral and lifestyle recommendations to help people manage diabetes or lose weight.

What unifies all of these companies is that they are leveraging the power of data and a growing understanding of human biology to change the world. What additionally unites them is that they all have servers overflowing with data - data that has troves of truly valuable knowledge locked inside. You, as a health data scientist, can significantly impacts the lives of thousands, even millions of people, by discovering actionable insights within this data, ultimately leading to better healthcare products, improve patient outcomes, or discover a novel treatment for a specific disease.

Your career will not only be financially rewarding and involve working on interesting problems with exceptional people, but can impact millions of people in a positive way. Being a great health data scientist at these world-leading organizations is not easy, and to do it effectively will take hard work and dedication. However, by leveraging cutting edge of technology, understanding the healthcare system, and working at the forefront of health data science, you can help create a world where healthcare is affordable and most importantly, effective.

# **Applying to Insight Health Data**

Applications are currently open for both Boston and Silicon Valley. We expect high demand based on previous interest, so we encourage you to apply early. Admission is granted on rolling basis.

To apply, please visit:

#### http://insighthealthdata.com/apply.html

**Please apply for the session that is in the geographic area where you would like to work after completing the program.** Interaction with mentor companies and Insight alumni is one of the most important parts of Insight. In the San Francisco session, you will network with companies all over Silicon Valley. In the Boston session, you will meet with companies based in the Great Boston area.

If you have any questions please email us at info@insighthealthdata.com.