# Project 1

# Pass NYC- Student identification and Recommendation for SHSAT

PASSNYC and its partners provide outreach services that improve the chances of students taking the SHSAT and receiving placements in these specialized high schools. The current process of identifying schools is effective, but PASSNYC could have an even greater impact with a more informed, granular approach to quantifying the potential for outreach at a given school. Proxies that have been good indicators of these types of schools include data on English Language Learners, Students with Disabilities, Students on Free/Reduced Lunch, and Students with Temporary Housing.

Part of this challenge is to assess the needs of students by using publicly available data to quantify the challenges they face in taking the SHSAT. The best solutions will enable PASSNYC to identify the schools where minority and underserved students stand to gain the most from services like after school programs, test preparation, mentoring, or resources for parents.

<https://www.kaggle.com/passnyc/data-science-for-good>

# Project 2

# Project with a heart( Source : Challenge in Datadriven

<https://www.drivendata.org/competitions/54/machine-learning-with-a-heart/page/109/>

Sample Data(180 values)



Goal : to predict whether there is a heart disease or not

# Project 3

# California Distributed Generation Interconnection Data Analysis

**Description:**

As above, PV energy is a hot topic amongst the STEM research community. However, instead of just looking at the output data generated from the United States as a whole, we’ll be focusing on the data collected through the all the interconnected solar systems within PG&E, SCE and SDG&E service territories within California. From these data sets, we have not only the power output but also a variety of other variables collected. From here, we can investigate the relationships between the utility associated with the PV’s system investor, the county the PV system is installed in, the type of mounting method, the cost of the project, the dependency of the project on whether or not there’s an electric vehicle associated with the system and more.

**Data Source:**

<https://www.californiadgstats.ca.gov/downloads/#_csi_15_id>

**Independent variables:**

* <https://www.californiadgstats.ca.gov/downloads/#_csi_15_id>
* <https://www.californiadgstats.ca.gov/static/documents/interconnection_nem_pv/NEM_CurrentlyInterconnectedDataKey.b29667e204f3.xlsx>