

**Data Driven Decision Making**

*Topic*

*Authors:*

Hunter Berberich

Brandon Botzer

Allison Gray

Victoria Ritorto

Ray Zaremski

**School of Graduate Professional Studies**

MPS/MS in Data Analytics

MPS/MS in Artificial Intelligence

DAAN 881 – Data Driven Decision Making

Summer, 2024

# Document Control

## Work carried out by:

*(This section is only required if the project is being completed by a group)*

|  |  |  |
| --- | --- | --- |
| **Name** | **Email Address** | **Task description** |
|  |  |  |
|  |  |  |
|  |  |  |

## Revision Sheet

|  |  |  |
| --- | --- | --- |
| **Release No.** | **Date** | **Revision Description** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**TABLE OF CONTENTS**

**General Guidelines**

1. Please use this template document to complete each deliverable assignment.
2. Each assignment must be submitted by the due date in the Course Schedule.
3. All figures should be followed by a brief description about the figure.
4. Figures can be hand-drawn and scanned in some circumstances, but the hand-drawn figure should be clear and legible to obtain full credit. Unclear hand-drawn figures will receive partial credit. For constructing figures and diagrams it is advised to use tools.
5. Figures and tables should have appropriate captions. For documenting and referencing styles please follow the APA or MLA writing style.
6. Please make sure that you provide a reference section.
7. Any material text or figure taken from books, journals or Internet should be referenced. If you have a sentence or a figure that does not belong (authorship) to you, they must be clearly referenced. If you fail to do so your report will be considered as a case for plagiarism. It is your duty to make sure that your report is free from any activity related to plagiarism. Please see the section on Academic Integrity found below. The penalty for plagiarism will be a “0” awarded to your report. Thus, it is good to keep it simple, always have the principle to acknowledge people for their contributions.

#### **Academic Integrity**

Academic integrity — scholarship free of fraud and deception — is an important educational objective of Penn State. Academic dishonesty can lead to a failing grade or referral to the [Office of Student Conduct](http://www.sa.psu.edu/ja/).

Academic dishonesty includes, but is not limited to:

* cheating
* plagiarism
* fabrication of information or citations
* facilitating acts of academic dishonesty by others
* unauthorized prior possession of examinations
* submitting the work of another person or work previously used without informing the instructor and securing written approval

In cases where academic integrity is questioned, [procedure requires an instructor to notify a student](http://www.psu.edu/oue/aappm/G-9-academic-integrity.html) of suspected dishonesty before filing a charge and recommended sanction with the college. Procedures allow a student to accept or contest a charge. If a student chooses to contest a charge, the case will then be managed by the respective college or campus Academic Integrity Committee. If a disciplinary sanction also is recommended, the case will be referred to the [Office of Student Conduct](http://www.sa.psu.edu/ja/title=).

All Penn State colleges abide by this policy, but review procedures may vary by college when academic dishonesty is suspected. Information about Penn State's academic integrity policy and college review procedures is included in the information that students receive upon enrolling in a course.

Penn State students are expected to act with civility and personal integrity; respect other students' dignity, rights, property, and help create and maintain an environment in which all can succeed through the fruits of their own efforts. An environment of academic integrity is requisite to respect for oneself and others, and a civil community.

#### For More Information on Academic Integrity at Penn State

Please see the [Academic Integrity Chart](http://www.campuses.psu.edu/CAO.pdf)  for specific college contact information or visit one of the following URLs:

* Penn State Senate [Policy on Academic Integrity](http://www.psu.edu/dept/oue/aappm/G-9.html)
* [iStudy for Success!](http://istudy.psu.edu/tutorials/) — learn about plagiarism, copyright, and academic integrity through an educational module
* [Turnitin](http://tlt.its.psu.edu/turnitin) a web-based plagiarism detection and prevention system

For **Deliverable 1**, your goal is to define a research or a business goal, identify the queries to be addressed, and identify the source(s) for collecting data. Please remember that data can be collected from single or multiple sources. In this week you must just identify the source(s) from where you are planning to collect data related to your research goals. The actual data collection steps can be conducted and completed in the following week. If you didn't read in the first week, now is the time to (again optional, but strongly recommended): Read the first three chapters of the book  [Analytical Skills for AI and Data Science (oreilly.com)](https://learning.oreilly.com/library/view/analytical-skills-for/9781492060932/) (you must be on Penn State network to read this, if you have issues accessing it get it through the e-books on PSU library website or reach out to me) to help you better understand how to define a business goal, identify the queries to be addressed.

**Deliverable 1:**Your team will come up with a business question and identify a problem or area for improvement based on the data source that you identify.  Your proposal should clearly identify your goal and how you will approach it by applying the topics presented in this course.

The project proposal should include:

* Problem Statement (the core problem or unmet need that your team will aim to solve)
* Possible resources available
* Potential Timeline
* Feasibility Estimate (*is this viable? do we have the resources?)*
* Anticipated Results or Deliverables

Be sure your proposal addresses the following:

***WHO?***

Start with the WHO: We should serve the community. Therefore, it is important to understand who are you trying to serve? Is it just this class? A customer? A company? Clearly identify this primary audience, then ask what "pain" they are experiencing that you can resolve? What do they want that they are not getting?

***WHAT?***

Once you accurately understand the WHO, you move to the WHAT: WHAT can you offer that will meet the need of who you are trying to serve? This is the critical question, because too often the question is asked the other way around: here is our WHAT, now WHO will buy it? Leading with our products or services and believing that the market will follow, risks irrelevance. By choosing who you serve first, understanding their needs, then framing your offering in light of these needs, you will be considered seriously--seeing that you "get" your customers.

***HOW?***

Once your WHAT is defined to be the true remedy to the unmet need of your WHO, you must determine HOW to provide the solution. The HOW involves considering the technological aspects of the project. You will need to consider how the topics we cover in the course are relevant to your project and how you might be able to integrate them into your solution. How might these topics help to transform your future, help improve businesses, solve problems, or provide an unmet need?

***TIMELINE***

When you finish defining who, what and how, please make and share a plan how you will work on your team project. Some of the details of the plan are already given as you will explore the topics in weekly manner. However, you should think of and make a plan when you are going to incorporate all the pieces into a story (write-up for the final report) which should contain introduction, related work, approach, experimental section with results, discussion, and conclusion.

**A similar grid as presented in *Analytical Skills for AI and Data Science:***

|  |  |  |
| --- | --- | --- |
| **Problem:** | **Whys** | **Value Potential** |
| **Initial Question:** |  |  |
| **Descriptive Question:** |  |  |
| **Predictive Question:** |  |  |
| **Prescriptive Question:** |  |  |

**Possible Datasets related to Opioids:**

<https://www.kaggle.com/datasets/jazzang/opioid-overdose-deaths/data>

<https://www.kaggle.com/datasets/apryor6/us-opiate-prescriptions> [how often providers perscribed]

<https://www.kaggle.com/datasets/craigchilvers/opioids-in-the-us-cdc-drug-overdose-deaths>

<https://www.kaggle.com/datasets/craigchilvers/opioids-vssr-provisional-drug-overdose-statistics>

<https://www.kaggle.com/datasets/ryanandreweckberg/opioid-crisis-by-interpersonal-relationships>

<https://www.kaggle.com/datasets/thedevastator/us-opioid-overdose-deaths>

<https://www.kaggle.com/datasets/programmerrdai/cocaineopioidscannabis-and-illicit-drugs>