Problem Statement: Employment Based Immigration

America is a melting pot. The American Dream is a beacon that draws the driven and the talented from all over the world. For entrepreneurs, engineers and educators alike, the path to the American Dream goes through one of the employment based immigration processes that lead to a permanent residency – aka the "green card". Employment based immigration has been vital to our competitiveness and in the coming age of AI, it will likely remain crucial for our future prosperity. Consider that the landmark 2017 paper (Titled: Attention is All You Need) that launched the current Generative AI revolution was written by a team comprised of six recent immigrants and two children of immigrants. For Datathon 2025, you are asked to explore and analyze past 10 years of data on employment based immigration to the USA.

Employment Based Immigration

First a quick primer to the Employment Based Immigration process. You are urged to look at the Dept of State Website on Employee Based Immigration for further details.

All Employment Based Immigration petitions, except for E1 and E5 categories, need a **Permanent Labor Certification** (often called **PERM**, which, very cleverly, stands for Program Electronic Review Management) from the Dept of Labor. Once PERM process is approved for an applicant, then a form I-140 Immigrant Petition for Alien Workers can be filed with the USCIS (U.S. Citizenship and Immigration Services) by the employer. Upon approval of I-140, the applicant and any dependents get in the line to get a Permanent Residency Card (aka the Green Card) when one becomes available. There is often a several years delay between getting approved for a green card and actually getting a green card since number of green cards issued is controlled based on the country of birth of the applicant (see Visa Bulletin issued by the Dept of State for current wait times for different Green Card categories and different countries of birth for the applicants).

For Michigan Ross Datathon 2025, you are provided with data on all PERM applications processed by the US Dept of State for the past 10 years. You are asked to explore and analyze this data to answer relevant business questions.

Your Task

Your goal is to analyze the provided datasets (described in detail below), potentially in combination with supplementary public information, to ask relevant business questions and answer them. Please note that we are not providing specific question prompts by design. You are asked to explore what questions can be asked and answered using the provided data. Here we are taking a broad view of "business" and who the potential audience for your analysis might be.

You are asked to pose your own question, explore your own solution approach, and provide answers/solutions using the available datasets in the available time. What is important is the insightfulness and depth of your conclusions and analysis. You need not be comprehensive; quality data analysis will be rewarded over the breadth of the question posed. You however must look at your task as broader than just data analysis – **you should look for a business question and consider the business context**. Note that thoughtfully prepared data visualizations are an integral part of your data exploration and analysis.

Creativity in formulating your own well-thought-out question or analysis approach generally has a positive effect on judges' assessment of your submission; however, it should not be at the expense of analytical depth, precision, and rigor, which are far more important. Please note that the provided datasets contain a lot of information – and you need not use all or even most of it for your chosen business question. You may decide to focus on a specific narrow slice/dice of the dataset that truly speaks to you and go deeper in your chosen narrow focus.

Your analysis should be well communicated to a non-technical audience, potentially with the help of appropriate visualizations.

Datasets

Datasets are provided in CSV format. CSV files can be easily opened with all data analysis tools including MS Excel. Please note that the datasets are **somewhat cleaned** while keeping much of the original variables available from the Dept. of Labor sources, therefore you may need to decide what factors to use before you can proceed to data analysis.

The datasets are sourced from the Office of Foreign Labor Certification (OFLC). This is the Dept. of Labor agency that conducts the PERM process.

You are provided with the PERM processing disclosure datasets provided by the US Dept of Labor OFLC for the past 10 years. For each year, you have following information (or columns in the CSV file) provided in the data dictionary in the next section.

Caution: you must first spend enough time to understand the dataset. Datathon Canvas discussion forums are a great place to ask and answer questions about the data. **Please note that you need not use every bit of data provided – use the part of the data that speaks to you, that sparks curiosity in you, that lends insights to you.**

Additional Datasets

You are welcome to look for additional datasets/information to supplement your analysis. All additional data used should be public and their source must be documented in the submission. Even though you can use additional dataset to supplement the provided dataset, you must keep the provided dataset central to your analysis. While evaluation of your submission will take a holistic approach and evaluate your entire submission, we will consider whether you have done justice to the provided dataset or not.

Data Dictionary

The table below lists all the columns available in the provided dataset and a description of what the columns contain. Please note that the data dictionary is our interpretation of the available data and might not be exhaustive in its coverage. Further, some data elements might not be available for all years.

While the datasets provided are in the form of yearly files (like one for 2024, one for 2023 and so on) – please note that the years indicate fiscal years for the DoL OLFC which goes from Oct to Sep. So a file for year 2024 has data from Sep 2023 to Oct 2024.

FIELD NAME	DESCRIPTION
CASE_NUMBER	Unique identifier assigned to each application submitted for processing to OFLC.
CASE_STATUS	Status associated with the last significant event or decision. Valid values include "Certified", "Certified-Expired", "Denied", and "Withdrawn".
RECEIVED_DATE	Date the application was received by OFLC.
DECISION_DATE	Date on which the determination was issued by OFLC
REFILE	Identifies if application was previously filed. Y = Employer has previously submitted an Application. N = Employer has not previously submitted an Application. Form ETA-9089, Section A, Item 1.
ORIG_FILE_DATE	Date application was originally filed (if applicable). Form ETA-9089, Section A, Item 1-A.
PREVIOUS_SWA_CASE_NUMBER_STATE	The SWA Case Number or State of the original filed application (if applicable). Valid Values will include a SWA Case Number or a State. Form ETA-9089, Section A, Item 1-B.
SCHD_A_SHEEPHERDER	Identifies whether the application is in support of Schedule A or a Sheepherder Occupation. Y = Application is in support of a Schedule A or Sheepherder Occupation. N = Application is not in support of a Schedule A or Sheepherder Occupation. Form ETA-9089, Section B, Item 1.

EMPLOYER_NAME	Legal business name of employer requesting permanent labor certification. Form ETA-9089, Section C, Item 1.
EMPLOYER_ADDRESS_1	Contact information of the employer requesting permanent labor certification. Form ETA-9089, Section C, Items 2 through 4.
EMPLOYER_ADDRESS_2	
EMPLOYER_CITY	
EMPLOYER_STATE_PROVINCE	
EMPLOYER_POSTAL_CODE	
EMPLOYER_NUM_EMPLOYEES	Total Number of employees employed by employer. Form ETA-9089, Section C, Item 5.
EMPLOYER_YEAR_COMMENCED_BUSINESS	Year the employer commenced business or incorporated. If the employer is a private household employing a household domestic worker, this question may be skipped. Form ETA-9089, Section C, Item 6.
EMPLOYER_FEIN	FEIN of the employer requesting permanent labor certification. Form 9089, Section C, Item 7.
NAICS_CODE	Industry code associated with the employer requesting permanent labor certification, as classified by the North American Industrial Classification System (NAICS). Form ETA-9089, Section C, Item 8.
FW_OWNERSHIP_INTEREST	Identifies if the foreign worker has ownership interest or familial relationship with the Employer. Y = Foreign Worker has Ownership Interest; N = Foreign Worker has no Ownership Interest. Form ETA9089, Section C, Item 9.
AGENT_ATTORNEY_NAME	Name of the agent or attorney designated to act on behalf of the employer for this application. Form ETA-9089, Section E, Item 1.
AGENT_ATTORNEY_FIRM_NAME	Name of the law firm or business filing a PERM application on behalf of the employer. Form ETA-9089, Section E, Item 2.
AGENT_ATTORNEY_ADDRESS_1	Contact information of the Agent or Attorney representing the Employer requesting permanent labor certification. Fields are located on the Form ETA-9089, Section E, Items 4 through 6.
AGENT_ATTORNEY_ADDRESS_2	
AGENT_ATTORNEY_CITY	
AGENT_ATTORNEY_STATE_PROVINCE	
AGENT_ATTORNEY_COUNTRY	
AGENT_ATTORNEY_POSTAL_CODE	

PW_TRACK_NUMBER	Unique identifier assigned to the Prevailing Wage Determination associated with the job opportunity. Form ETA-9089, Section F, Item 1
PW_SOC_CODE	Occupational code associated with the job being requested for permanent labor certification, as classified by the Standard Occupational Classification (SOC) System. Form ETA-9089, Section F, Item 2.
PW_SOC_TITLE	Occupational title associated with the SOC/O*NET Code. Form ETA9089, Section F, Item 3
PW_SKILL_LEVEL	Level of the prevailing wage determination. Valid values include "Level I", "Level II", "Level III", "Level IV", and "N/A". Form ETA-9089, Section F, Item 4.
PW_WAGE	Prevailing wage for the job being requested for permanent labor certification. Form ETA-9089, Section F, Item 5.
PW_UNIT_OF_PAY	Unit of Pay. Valid values include "Hour", "Week", "Bi-Weekly", "Month", and "Year". Form ETA-9089, Section F, Item 5
PW_WAGE_SOURCE	Name of the entity providing prevailing wage information for the job. Valid prevailing wage source codes are: "OES"; "CBA"; "Employer Conducted Survey"; "DBA"; "SCA"; and "Other". Form ETA-9089, Section F, Item 6.
PW_SOURCE_NAME_OTHER	Name of the other source of prevailing wage information for the job. Form ETA-9089, Section F, Item 6-A.
PW_DETERMINATION_DATE	Date on which the Prevailing Wage decision was issued. Form ETA9089, Section F, Item 7.
PW_EXPIRATION_DATE	Date on which the Prevailing Wage will expire. Form ETA-9089, Section F, Item 8
WAGE_OFFER_FROM WAGE_OFFER_TO	Employer's offered wage. Form ETA-9089, Section G, Item 1
WAGE_OFFER_UNIT_OF_PAY	Unit of Pay. Valid values include "Hour", "Week," "Bi-Weekly," "Month", and "Year". Form ETA-9089, Section G, Item 1.
WORKSITE_ADDRESS_1 WORKSITE_ADDRESS_2 WORKSITE_CITY WORKSITE_STATE WORKSITE_POSTAL_CODE	Primary worksite address or location. Fields are located on the Form ETA-9089, Section H, Items 1 and 2.

JOB_TITLE	Title of the permanent job. Form ETA-9089, Section H, Item 3
MINIMUM_EDUCATION	The minimum U.S. diploma or degree required by the employer for the position. Valid values include "None", "High School", "Associate's", "Bachelor's", "Master's", "Doctorate, and "Other". Form ETA-9089, Section H, Item 4
JOB_EDUCATION_MIN_OTHER	If "Other", the specific U.S. diploma or degree required. Form ETA-9089, Section H, Item 4-A
MAJOR_FIELD_OF_STUDY	Major or field of study required by the employer for the job opportunity. Form ETA-9089, Section H, Item 4-B
REQUIRED_TRAINING	Identifies whether or not training is required for the job. Y = Training is required for the job. N = Training is not required for the job. Form ETA9089, Section H, Item 5
REQUIRED_TRAINING_MONTHS	If training is required, number of months training. Form ETA-9089, Section H, Item 5-A
REQUIRED_TRAINING_FIELD	If training is required, field of training. Form ETA-9089, Section H, Item 5-B.
REQUIRED_EXPERIENCE	Identifies whether experience in the job offered is a requirement. Y = Experience in the job is required. N = Experience in the job is not required. Form ETA-9089, Section H, Item 6
REQUIRED_EXPERIECE_MONTHS	If experience in the job offered is required, the number of months required. Form ETA-9089, Section H, Item 6-A.
ACCEPT_ALT_FIELD_OF_STUDY	Indicates if an alternate field of study will be acceptable for education requirement. Y = Alternate field of study is acceptable. N = Alternate field of study is not acceptable. Form ETA-9089, Section H, Item 7
ACCEPT_ALT__MAJOR_FLD_OF_STUDY	Alternate field of study that will be acceptable for education requirement (if applicable). Form ETA-9089, Section H, Item 7-A.
ACCEPT_ALT_COMBO	Indicates if an alternate combination of education and experience will be acceptable. Y = An alternate combination of education and experience will be acceptable. N = An alternate combination of education and experience will not be acceptable. Form ETA-9089, Section H, Item 8.
ACCEPT_ALT_COMBO_EDUCATION	If applicable, the alternate level of education that is required. Valid values include "None", "High School", "Associate's", "Bachelor's", "Master's", "Doctorate, and "Other".. Form ETA-9089, Section H, Item 8- A.

ACCEPT_ALT_COMBO_ED_OTHER	Indicates the "other" alternate level of education that is required. Form ETA-9089, Section H, Item 8-B.
ACCEPT_ALT_COMBO_EDUCATION_YRS	The number of years of experience that is acceptable. Form ETA-9089, Section H, Item 8-C.
ACCEPT_FOREIGN_EDUCATION	Indicates if a foreign educational equivalent is acceptable. Y = Foreign educational equivalent is acceptable. N = Foreign educational experience is not acceptable. Form ETA-9089, Section H, Item 9
ACCEPT_ALT_OCCUPATION	Indicates if experience in an alternate occupation is acceptable. Y = Experience is an alternate occupation is acceptable. N = Experience in an alternate occupation is not acceptable. Form ETA-9089, Section H, Item 10.
ACCEPT_ALT_OCCUPATION_MONTHS	The number of months of experience in the alternate occupation (if applicable). Form ETA-9089, Section H, Item 10-A.
ACCEPT_ALT_JOB_TITLE	The job title of the alternate occupation that is acceptable (if applicable). Form ETA-9089, Section H, Item 10-B.
JOB_OPP_REQUIREMENTS_NORMAL	Indicates if the job opportunity's requirements are normal for the occupation being offered. Y = Job opportunity's requirements are normal for the occupation being offered. N = Job opportunity's requirements are not normal for the occupation being offered. Form ETA-9089, Section H, Item 12.
FOREIGN_LANGUAGE_REQUIRED	Indicates if knowledge of a foreign language is required to perform the job duties. Y = Knowledge of a foreign language is required to perform the job duties. N = Knowledge of a foreign language is not required to perform the job duties. Form ETA-9089, Section H, Item 13
SPECIFIC_SKILLS	Specific skills or job related requirements for the job opportunity. Form ETA-9089, Section H, Item 14.
COMBINATION_OCCUPATIONN	Identifies whether or not the job includes a combination of occupations. Y = Job includes a combination of occupations. N = Job does not include a combination of occupations. Form ETA-9089, Section H, Item 15.
OFFERED_TO_APPL_FOREIGN_WORKER	Indicates if the position identified in this application is being offered to the foreign worker being sponsored on

	the application. Y = Position identified in this application is being offered to the foreign worker being sponsored on this application. N = Position identified in this application is not being offered to the foreign worker being sponsored on this application. Form ETA-9089, Section H, Item 16
FOREIGN_WORKER_LIVE_ON_PERM	Indicates if the job requires the foreign worker to live on the employer's premises. Y = Job requires the foreign worker to live on the employer's premises. N = Job does not require the foreign worker to live on the employer's premises. Form ETA-9089, Section H, Item 17.
FOREIGN_WORKER_LIVE_IN_DOM_SER	Identify whether the application is for a live-in domestic service worker. Domestic service workers refer to "private household workers." Y = Application is for a live-in domestic service worker. N = Application is not for a live-in domestic service worker. Form ETA-9089, Section H, Item 18
FOREIGN_WORKER_LIVE_IN_DOM_SVC_CNT	Identifies whether the employer and the foreign worker have executed an employment contract and the employer has provided a copy of the contract to the foreign worker. Y = The employer and the foreign worker have executed an employment contract and the employer has provided a copy of the contract to the foreign worker. N = The employer and the foreign worker have not executed an employment contract and/or the employer has not provided a copy of the contract to the foreign worker. N/A = Not applicable. Form ETA-9089, Section H, Item 18-A
PROFESSIONAL_OCCUPATION	Identifies whether or not the application is for a professional occupation, other than a college or university professor. Y = Application is for a professional occupation other than a college or university professor. N = Application is not for a professional occupation other than a college or university professor. Form ETA-9089, Section I.a, Item 1.
APP_FOR_COLLEGE_U_TEACHER	Identifies whether or not the application is for a college or university teacher. Y = Application is for a college or university professor. N = Application is not for a college or university professor. Form ETA9089, Section I.a, Item 2

COMPETITIVE_PROCESS	If application is for a College or University Teacher, Y = Worker was selected using a competitive recruitment and selection process (if applicable); N = Worker was not selected using a competitive recruitment and selection process (if applicable). If blank, application is not for a College or University Teacher. Form ETA-9089, Section I.a, Item 2-A
BASIC_RECRUITMENT_PROCESS	If application is for a College or University Teacher. Y = Worker was selected using a basic recruitment process for professional occupations; N = Worker was not selected using a basic recruitment process for professional occupations. Field is located on the Form ETA-9089, Section I.a, Item 2-B.
TEACHER_SELECT_DATE	The date the foreign worker was selected using the competitive recruitment and selection process. Form ETA-9089, Section I.b, Item 3.
TEACHER_PUB_JOURNAL_NAME	The name and date of the national professional journal in which the advertisement was placed. Form ETA-9089, Section I.b, Item 4.
SWA_JOB_ORDER_START_DATE	The start date for the State Workforce Agency job order. Form ETA9089, Section I.c, Item 6.
SWA_JOB_ORDER_END_DATE	The end date for the State Workforce Agency job order. Form ETA9089, Section I.c, Item 7.
SUNDAY_EDITION_NEWSPAPER	Indicates if there is a Sunday edition of the newspaper in the area of intended employment. Y = There is a Sunday edition of the newspaper in the area of intended employment. N = There is not a Sunday edition of the newspaper in the area of intended employment. Form ETA-9089, Section I.c, Item 8.
FIRST_NEWSPAPER_NAME	The name of the newspaper (of general circulation) in which the first advertisement was placed. Form ETA-9089, Section I.c, Item 9.
FIRST_ADVERTISEMENT_START_DATE	The date the first advertisement was published. Form ETA-9089, Section I.c, Item 10.
SECOND_NEWSPAPER_AD_NAME	The name of the newspaper or professional journal in which the second advertisement was placed (if applicable). Form ETA-9089, Section I.c, Item 11.
SECOND_ADVERTISEMENNT_TYPE	Indicates whether the advertisement ran in a Newspaper or Journal. Form ETA-9089, Section I.c, Item 11.

SECOND_AD_START_DATE	The date of the second newspaper advertisement or date of publication of journal (if applicable). Form ETA-9089, Section I.c, Item 12
JOB_FAIR_FROM_DATE JOB_FAIR_TO_DATE	Dates advertised at job fair (if applicable). Form ETA-9089, Section I.d, Item 13
ON_CAMPUS_RECRUITING_FROM_DATE ON_CAMPUS_RECRUITING_TO_DATE	Dates of on-campus recruiting (if applicable). Form ETA-9089, Section I.d, Item 14
EMPLOYER_WEBSITE_FROM_DATE EMPLOYER_WEBSITE_TO_DATE	Dates advertised on the employer's website (if applicable). Form ETA 9089, Section I.d, Item 15.
PRO_ORG_AD_FROM_DATE PRO_ORG_ADVERTISEMENT_TO_DATE	Dates advertised with a trade or professional organization (if applicable). Form ETA-9089, Section I.d, Item 16
JOB_SEARCH_WEBSITE_FROM_DATE JOB_SEARCH_WEBSITE_TO_DATE	Dates listed with a job search website (if applicable).
PVT_EMPLOYMENT_FIRM_FROM_DATE PVT_EMPLOYMENT_FIRM_TO_DATE	Dates listed with a private employment firm (if applicable)
EMPLOYEE_REF_PROG_FROM_DATE EMPLOYEE_REFERRAL_PROG_TO_DATE	Dates advertised with an employee referral program (if applicable)
CAMPUS_PLACEMENT_FROM_DATE CAMPUS_PLACEMENT_TO_DATE	Dates advertised with a campus placement office (if applicable)
LOCAL_ETHNIC_PAPER_FROM_DATE LOCAL_ETHNIC_PAPER_TO_DATE	Dates advertised with a local or ethnic newspaper (if applicable)
RADIO_TV_AD_FROM_DATE RADIO_TV_AD_TO_DATE	Dates advertised with radio or TV stations (if applicable).
EMP_RECEIVED_PAYMENT	Indicates if the employer received payment of any kind for the submission of this application. Y = The employer received payment. N= The employer did not receive payment. Form ETA-9089 Section I.e, Item 23.
PAYMENT_DETAILS	If the employer received payment of any kind for the submission of this application, details of that payment including the amount, date, and purpose of that payment. Form ETA-9089, Section I.e, Item 23-A.
LAYOFF_IN_PAST_SIX_MONTHS	Indicates if the employer had a layoff in the occupation or related occupation in the area of intended employment within the past six months. Y = The employer had a layoff. N = The employer did not have a layoff

US_WORKERS_CONSIDERED	Indicates whether the laid off U.S. workers were notified and considered for the job opportunity for which certification is sought. Y = U.S. workers were notified and considered for the job opportunity for which certification is sought. N = U.S. workers were not notified and considered for the job opportunity for which certification is sought. N/A = Not applicable
COUNTRY_OF_CITIZENSHIP	Country of citizenship of the foreign worker being sponsored by the employer for permanent employment in the United States
FOREIGN_WORKER_BIRTH_COUNTRY	Foreign worker's country of birth
CLASS_OF_ADMISSION	Indicates the current visa status of the foreign worker.
FOREIGN_WORKER_EDUCATION	Highest Education achieved by the foreign worker. Valid values include "None", "High School", "Associate's", "Bachelor's", "Master's", "Doctorate", and "Other".
FOREIGN_WORKER_EDUCATION_OTHER	Identifies the "other" highest level of education achieved by the foreign worker.
FOREIGN_WORKER_INFO_MAJOR	Major field(s) of study in reference to the highest level achieved by the foreign worker.
FOREIGN_WORKER_YRS_ED_COMP	Year the relevant education was completed by the foreign worker
FOREIGN_WORKER_INST_OF_ED	Name of the institution where the relevant education was achieved by the foreign worker.
FOREIGN_WORKER_ED_INST_ADD_1	Geographic information for the educational institution identified on the Form
FOREIGN_WORKER_ED_INST_ADD_2	
FOREIGN_WORKER_ED_INST_CITY	
FOREIGN_WORKER_ED_INST_STATE_P	
FOREIGN_WORKER_ED_INST_COUNTRY	
FOREIGN_WORKER_ED_INST_POST_CD	
FOREIGN_WORKER_TRAINING_COMP	Indicates whether the foreign worker completed the training required for the requested job opportunity. Y = The foreign worker completed the training. N = The foreign worker did not complete the training. N/A = Not applicable.
FOREIGN_WORKER_REQ_EXPERIENCE	Indicates whether the foreign worker has the experience as required for the requested job opportunity. Y = The foreign worker has the experience. N = The foreign worker does not have the experience. N/A = Not applicable

FOREIGN_WORKER_ALT_ED_EXP	Indicates whether the foreign worker possesses the alternate combination of education and experience. Y = The foreign worker possesses the alternate combination of education and experience. N = The foreign worker does not possess the alternate combination of education and experience. N/A = Not applicable
FOREIGN_WORKER_ALT_OCC_EXP	Indicates whether the foreign worker has the experience in an alternate occupation as indicated in question H.10 of ETA Form 9089. Y = The foreign worker has the experience. N = The foreign worker does not have the experience. N/A = Not applicable.
FOREIGN_WORKER_EXP_WITH_EMPL	Indicates whether the foreign worker gained any of the qualifying experience with the employer in a position substantially comparable to the job opportunity requested. Y = The foreign worker gained qualifying experience with the employer. N = The foreign worker did not gain qualifying experience with the employer. N/A = Not applicable
FOREIGN_WORKER_EMPL_PAY_FOR_ED	Indicates whether the employer paid for any of the foreign worker's education or training necessary to satisfy any of the employer's job requirements for this position. Y = The employer paid for foreign worker's education. N = The employer did not pay for foreign worker's education.
FOREIGN_WORKER_CURR_EMPLOYED	Indicates whether the foreign worker is currently employed with the petitioning employer. Y = The foreign worker is currently employed with the petitioning employer. N = The foreign worker is not currently employed with the petitioning employer.
EMPLOYER_COMPLETED_APPLICATION	Indicates whether the application was completed by employer. Y = The application was completed by the employer. N = The application was not completed by the employer.
PREPARER_NAME	Name of person who is responsible for the preparation of this application.
PREPARER_TITLE	Occupational title of person that prepared application.
EMP_INFO_DECL_NAME	Name of person signing the Employer Declaration
EMP_DECL_TITLE	Title of person signing the Employer Declaration.

You would note that not all years contain all the information. You would specifically find that last five years have more information that the five years before that. Plan your data exploration and analysis accordingly.

Submission: Content

You are expected to submit two separate deliverables: a recorded presentation of max 8 min duration and a presentation file. Your submissions must be in easily accessible formats (e.g PPT/PDF for presentation file, mpeg for recorded video). You can use any tool like Zoom for creating your recorded presentation video.

Additional information that may not fit the presentation time may be placed in an appendix of the presentation file. You will be asked to make your submissions to a Canvas assignment. Note that deadlines are strictly enforced and any significant delay in submission will result in your submission being rejected.

Judges will be evaluating your work without your team there to explain it; therefore, your submission must "**speak for itself**". It need not be polished to the level of a final product, but do ensure that your main findings are clear and that any visualizations are functionally labeled and interpreted.

Note that since we don't provide a prompt, you are free to choose your audience. You may look at the data from the point of view of emissions, firms, customers, regulatory bodies, or any other stakeholders. However, you should clearly document your audience and the value you seek to provide to them.

You are asked to build and submit a presentation file - this is the file you should have used for your recorded presentation – and you will use to present your solution live to the expert panel of judges should you be chosen as one of the 6 finalist teams.

Tips & Recommendations

We recommend that you follow a low-hanging-fruit first approach. While it is tempting to go for big and bold ideas, remember that you have limited time and even essential tasks like data cleaning and joining tables are likely to take significant time. We strongly recommend that you take care of low, effort-high return tasks first before going for more in-depth analytics. In particular, effective data visualization is an efficient, high return, low(er) effort target to finish first. It is important that you do not get stuck in a complex technical workflow and lose track of not just the time but also your overall analysis objective. If something is taking too long, then it is okay to shelve it for the moment and move on to more easily executable tasks. Prioritize simple, easy, doable tasks first.

We recommend that your team not try to learn new tools if possible; instead, leverage your existing skills to extract as much insight from the data as you can. You will find that even basic tools like Excel and Tableau, if used well, can be quite effective in a time bound competition like this where speed is essential. Further, a good business narrative and interpretation from a simple analysis is likely more valuable than complex technical analysis that does not lend itself to (or does not leave time for) a compelling business insight.

Note that your submission will be evaluated for its overall communication, structure, and narrative not just on your technical work. **If your submission does not adequately communicate the story coming**

out from your technical work then all your work is essentially wasted. Your submission should present a compelling story and that story should be clear even with a cursory look at your submission.

We STRONGLY encourage you to start building up your final submission AT LEAST two to three hours before the submission deadline. In the past, many teams have spent a lot of time conducting great analyses, only to realize that they left almost no time for actually writing up and presenting their results. This cannot be stressed enough – quality data analysis that is incomplete or poorly presented will NOT win the competition.

Potential Analysis Approaches

You have a lot of data and no fixed decision problem to solve. In such situation it is recommended that you keep in mind the following:

1. Before you do any "analysis", you would want to explore the data – just get to know what you have, how are the variables distributed, whether significant cleaning is needed. A through exploration of the data will not only show you any data cleaning need, it will also point towards insights that can form the basis for later solutions.
2. Once you have a sense of the data, you would want to get the two important questions sorted:
    a. Who is your audience? Who are you doing your analysis for? For the dataset provided, your audience could you USCIS – the agency that processes applications, Employers that hire, Applicants who are looking for green cards, local government official who are looking at impact of immigration on local economy, policy makers that regular immigration, investors that might want to invest or not in a company and so on. Your analysis can't solve everything for everyone – so step 1 is to choose who is going to be your audience.
    b. What problem are you solving for your audience? Your solution must be rooted in a need. You must provide some value for your audience. Identify the problem you are addressing and through your analysis, show how you will provide value for your audience.
3. Now that you have an audience and a problem, you would want to tell an effective story that will connect with and persuade your audience to make the needed decisions. All your analysis is worthless if that does not move the needle for your audience – and people are persuaded by stories and visuals – not by numbers and regressions for example. Tell an effective story.

Ask for Help

Do not hesitate to ask for help. You will have access to a Canvas discussion board where you can ask questions and the Datathon organizers will respond as appropriate. Representatives from sponsor companies will be offering Office Hours where you can get their advice. Datathon organizing team will be available throughout the day in case you have any technical or logistical issues – just let us know and we will do our best to assist.

If you wish to ask a question privately (i.e. not in the public Canvas Discussion Thread) then please send an email to Sanjeev (sankum@umich.edu).

That's it. We hope that you will have a wonderful experience in the Michigan Ross Datathon 2025 and that it will prove a fun learning experience for you.