

Trends in T-Visa Application

Kehan Wang
Xiaoxin Gan
Yansong Bai

1. Introduction & Problem Statement

This project will look at T-Visas, a visa program for victims of human trafficking. There has been an increase in denials in T-Visa applications (there are 5,000 available per year). Never have more than 1,000 been granted in a year and the past few years have seen both an increase in applications and an increase in denials. There is a hypothesis that the increase in denials is related to procedural barriers that immigration is constructing to make it more difficult for applications to find success. The data will look at administrative appeal decisions (AAO) to examine these trends in terms of appeals and the result of these appeals. The goal is to understand appeal T-visa case trends within the decisions by analyzing the presence of specific keywords within the texts and to conduct topic modeling. In this project, we analyzed the following questions. How many cases were decided (separated by year); how many cases were dismissed (separated by year to demonstrate trends); how many cases were granted (indicated by language including grant, sustain, or withdraw and remanded back to the agency); what grounds were cases denied (separated by year to demonstrate trend); how has the trend changed over time; and what are the patterns indicating by the T-Visa Application.

2. Methodology & Algorithm

2.1 Data Sources

The data sources include AAO Decisions from the website and LEXIS PDF files from our client. All AAO Decisions are found on the website [here](#). We analyzed [non-precedent decisions](#). In order to search within the AAO decisions for Applications for T nonimmigrant status. We search them by two queries, which are “application for t nonimmigrant status” and “form I-914”. We choose to use the “Form I-914” query because it is the more accurate one to use. Within this search query, we see a range of decision types. We only need to download decisions with the heading **Application for T Nonimmigrant Decision**. We will need to exclude derivative (family member) applications from this set. Figure 1 is an example of the type.

Figure 1: Example of **Application for T Nonimmigrant Decision**.

[PDF] JUN 24 2020 - Application for T **Nonimmigrant Status**
www.uscis.gov/sites/default/files/err/D12%20-%20Application%20for%20T%20Nonimmigrant%20Status/D...
 Form 1-914, Application for T **Nonimmigrant Status** Non-Precedent Decision of the...Form 1-914, Application for T
Nonimmigrant Status (T application) ...

2.2 Data Collection

Then, the next step is to collect all data from the data mentioned above sources. For AAO websites, we need to get all links to all T-Visa applications. We first search the keyword “Form I-914” to get all related applications.

Figure 2: Search Page pattern

search.usa.gov/search?affiliate=uscis-ao&dc=18478&page=1&query=Form+I-914&search=Search&utf8=✓

On the search result page, we found the URL has a specific pattern, as Figure 2 shows. Hence, the task should be relatively easy — we need to loop through all the pages by varying the page number, crawl all the links of applications, and download them into a single directory by sending HTTP requests (request package). For the LEXIS data source, thanks to our client, we have had all PDFs downloaded. We need to save them into the same directory so that we can do the data extraction.

2.3 Data Extraction

In step 3, in order to extract information from these downloaded files, we first convert them into texts by applying packages like pdfminer, PyPDF2. After that, as two data sources have completely different formats, we need to design a data extraction plan separately. After a series of attempts, we use regular expression operations (re package) to specify a set of strings that match a specific pattern, then extract them and save them into a CSV file. As a result, we come up with two data sets that consist of different columns. You can observe what each column represents in Appendix A, and how we design the extraction plan for strategic questions in Appendix B.

2.4 Data Analysis

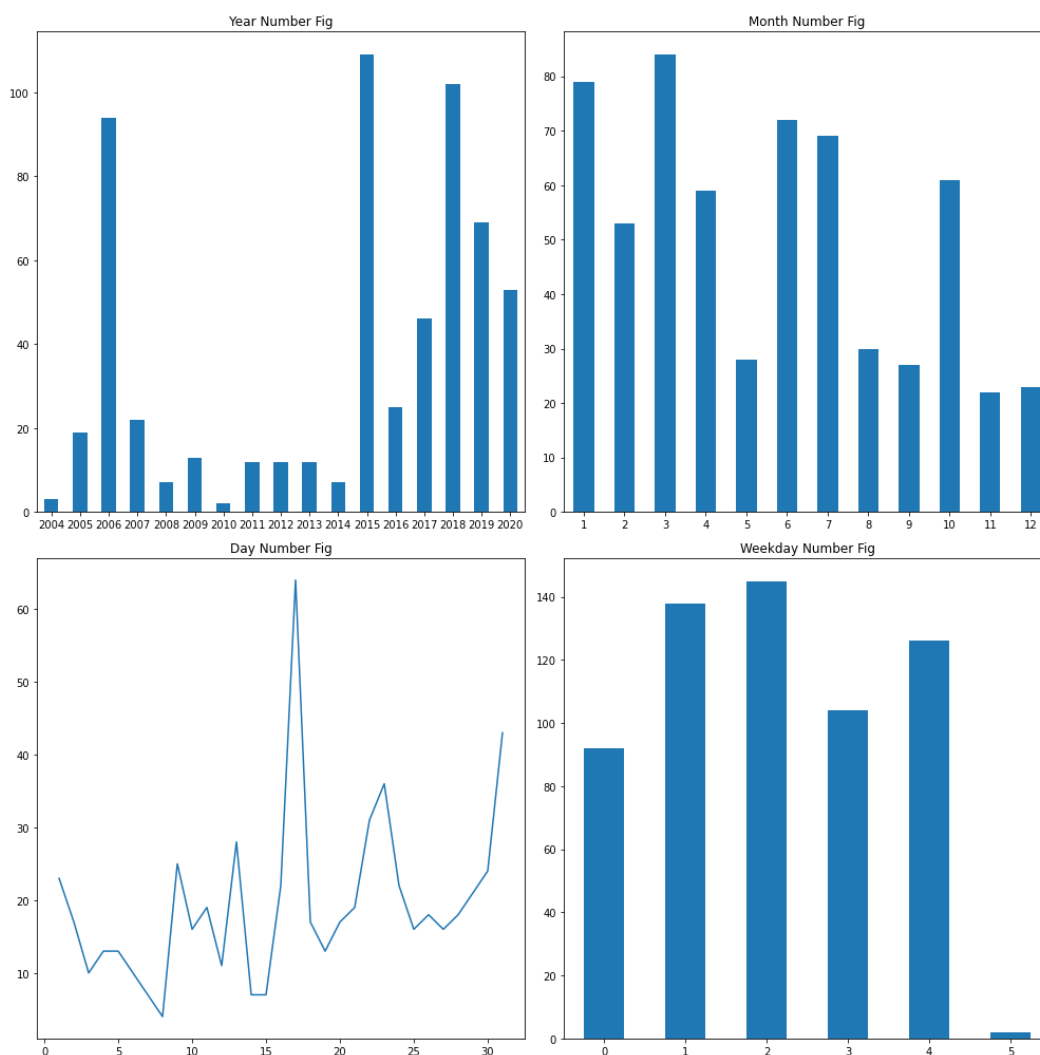
After getting all the data stored in CSV files, we want to find out what is the trend and what causes the difference. We concat two CSV files into one, and use matplotlib.pyplot to visualize the data by years, months, days, and weekdays where the trend can be easily observed. In order to find out what topics are mentioned most frequently in applications, we first did some analysis based on word count by using the CountVectorizer method. However, the method didn't offer us good results because there are too many meaningless words. So, we decided to use Latent Dirichlet Allocation to uncover hidden structure in the collection of texts as the package provides many good APIs, including dimensionality reduction, where results and patterns can be clearly evaluated.

3. Results & Observations

3.1 How many cases were decided?

There are a total of 188 cases that were decided from the AAO data source. Considering the LEXIS data sources, we found that there are a total of 502 cases. By concatenating two data sets and excluding derivative applications, there are a total of 630 “Appeal” cases are decided (occupied 86.67% of total cases). Our figure below has all the cases from 2004 to 2020; according to it, 2015 has the most applications, and the year 2010 has the fewest applications. We also observed that the middle of each month has the most number of applications, and Friday is not a good day to submit an application.

Figure 3: Trend for All “Appeal” Cases



3.2 How many cases were dismissed? (separated by year to demonstrate trends)

There are a total of 544 dismissed cases for the old dataset. According to the figure below, we can see that there are two peaks from 2004-2020. The first peak was in 2006. The cases dismissed in 2004-2006 and 2007-2014 are quite low. Besides that, we also noticed that the dismissed cases in 2015 and 2018 are quite high. In the rest of the years, the number of dismissed cases is at the average level. We conducted some work related to topic modeling. According to Figure 11, you can see the top 7 topics most likely to be mentioned in “dismissed” cases.

Figure 4: Trend for Case “Dismissed”

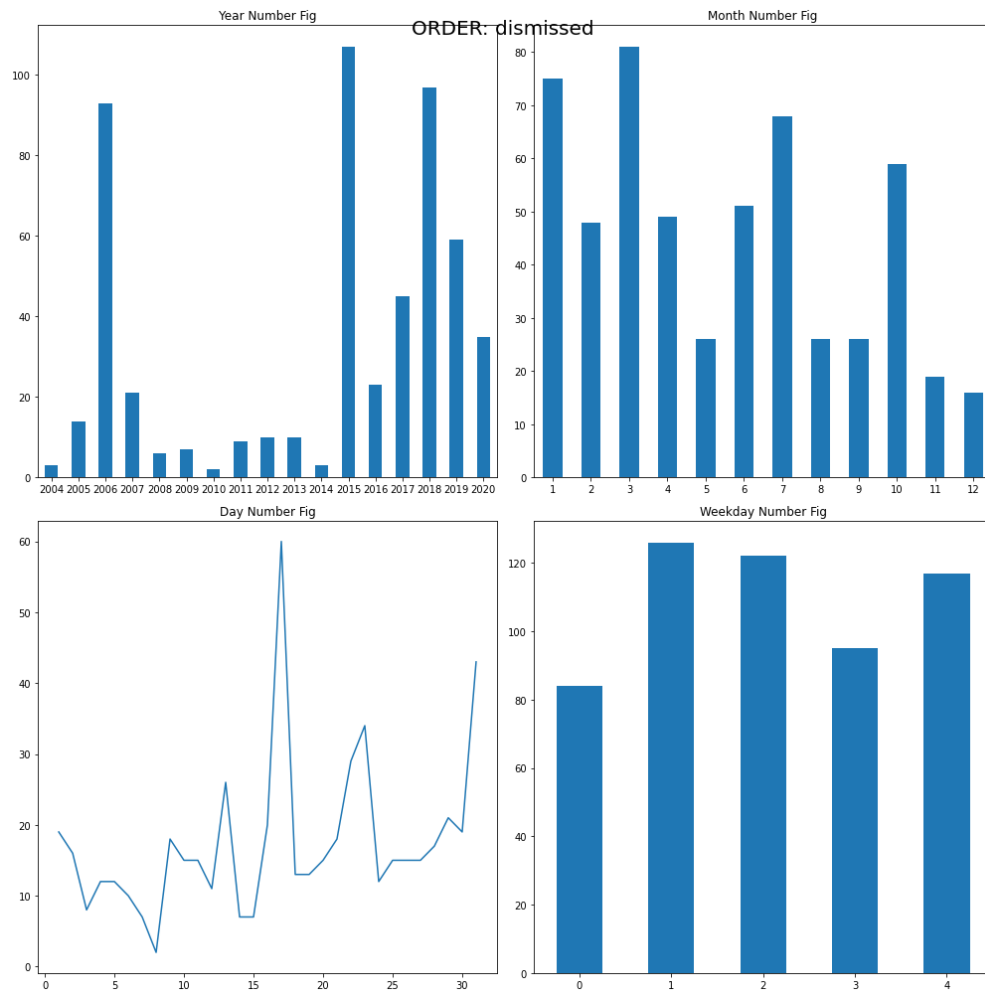


Figure 5: Top Topic for Case “Dismissed”

Topic #0:
 applicant application trafficking nonimmigrant form act appeal director victim 15 sections vermont center service status

Topic #1:
 applicant states united falcon pursuant act agreed employ order remain entered native citizen status immigration

Topic #2:
 trafficking act 15 applicant nonimmigrant classification victim severe seeks nationality 101 immigration persons form 1101

Topic #3:
 redacted applicant states united mr visas family act people money citizen student married work january

Topic #4:
 china native citizen immigration enter 2004 attempted judge september respondent motion 2005 smuggled continuance decision

Topic #5:
 trafficking severe act section form applicant crime persons investigation alien united states request account victim

Topic #6:
 aao waiver inadmissibility act issue united states section director rfe address traffickers immigrant october issued

3.3. How many cases were granted (this can be indicated by language, including grant, sustain, or withdrawal and remanded back to the agency)?

There are merely a total of 53 cases. The number of granted cases are in the low region compared to dismissed cases. The trend for granted cases is quite low from 2005-2018. Then, there is a big jump in the year 2019. We can see that the number of granted cases increases very quickly and strongly. The peak for granted cases is in the year 2020. We also conducted some work related to topic modeling in this part; according to Figure 13, we can top 7 topics most likely to be mentioned in “granted” cases.

Figure 6: Trend for the case “granted”

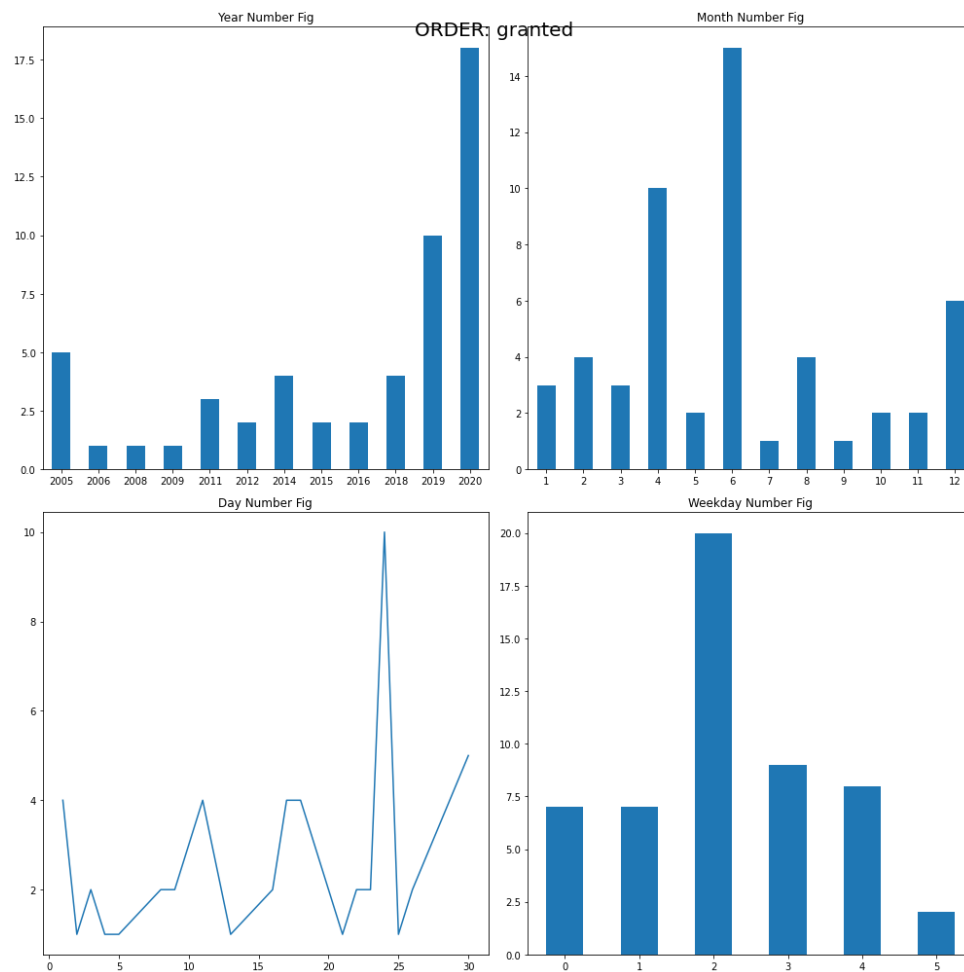


Figure 7: Top Topic for the case "granted"

Topic #0:
applicant application nonimmigrant trafficking form act appeal director victim 15 sections status 914 denied center

Topic #1:
position philippines states pay united order large debt 2003 act section stated remain placed immigration

Topic #2:
china 2004 attempted enter smuggled taken native august citizen custody fraudulent using passport country united

Topic #3:
trafficking applicant severe persons form victim act 15 states united director account present establish nonimmigrant

Topic #4:
act 15 applicant victim classification section form nonimmigrant severe nationality immigration 101 seeks trafficking persons

Topic #5:
nonimmigrant act classification trafficking 15 status seeks 101 immigration applicant acts nationality affords 1101 victims

Topic #6:
applicant states united falcon pursuant act agreed employ order citizen status native entered remain immigration

3.4. What grounds were cases denied? (separated by year to demonstrate trends)

We only found 1 denied case. The case was denied on December 13th, 2017. The core terms including traffic, physical presence, severe form, redact, nonimmigrant, abroad, the port of entry, sex act, classification, interim, sex, reasonable request, judicial process, time of filing, servitude, reenter, kill. These terms should be avoided in applications.

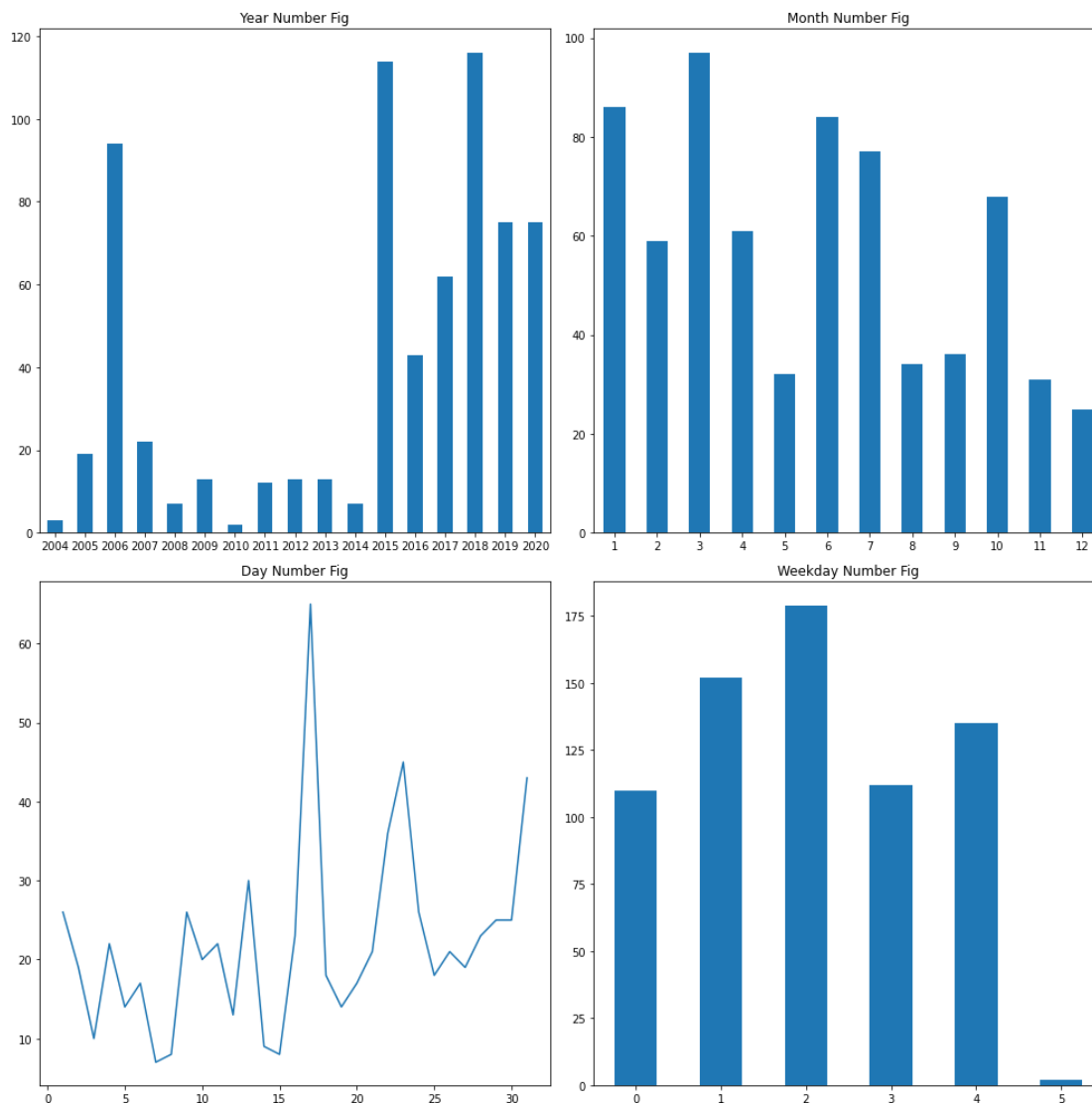
3.5. How has this changed over time?

The trend for the decided cases and dismissals we can see their trend looks quite similar, went up, went down, and then went up again. For granted cases, we can see that the number of cases stay relatively low for a long time and jump up erratically in recent years—the overall decided cases and granted cases all less than the number of dismissed cases.

3.6. Any patterns? (All cases, including appeal and motion)

Later, we plot a figure to find the trend of all applications from 2004 to 2020. Since motion cases are only a small portion of total cases, we didn't find any specific patterns that were different from Figure 9. However, according to Figure 10 and Figure 11, we can assume that June is a good month to apply. The applications submitted on the 25th of each month are most likely to be granted, and Tuesday is also a good day to start an application. Besides, during preprocessing, we found derivative applications (applications for family members) are most likely to be rejected/denied. Hence, derivative applications are not recommended.

Figure 8: Trend for all cases, including appeal and motion



Appendix A:

a. What does each column represent?

We made two datasets based on the given data sources, namely AAO Non-Precedent Decisions and LEXIS. We contacted them during the analysis stage.

'Url':	AAO Non-Precedent Decisions	LEXIS
'File_name'	This column is for the file name	

'ID'	This column is for the case number, a unique identifier to distinguish each case	
'date',	This column is for the date of the application	
'type'	This column is used to distinguish appeal cases or motion cases.	
'status'	Only apply to AAO Non-Precedent Decision, used to check if the applicant is the family member of victims.	Set to “Unknown” or “NaN”
'order'	The column is used to describe the result of the application, e.g dismissed or granted.	
'is_family'	Only apply to AAO Non-Precedent Decision, a boolean value to determine whether the applicant is a family member	Set to “Unknown” or “NaN”
'description'	This Column declares the description of the application.	
'core_terms'	Set to “Unknown” or “NaN”	Only apply to LEXIS cases, this column declares the description of the application.
'citation'	Set to “Unknown” or “NaN”	Only apply to LEXIS cases, this column is for the head of the application record in the LEXIS system. The four/five-digit number can be used to distinguish each application.
'S-M-P'	Set to “Unknown” or “NaN”	Only apply to LEXIS cases, this column describes whether the application contains “Matter of S-M-P”
'aao citation'	Set to “Unknown” or “NaN”	Only apply to LEXIS cases, this column describes the aao citation of each case.

'consel'	Set to “Unknown” or “NaN”	Only apply to LEXIS cases, this column is to illustrate the representation of the applicant.
'description'	This column contains a description of the application.	

Appendix B:

a. How many cases were decided? (separated by year)

This can be determined by downloading the decisions for “Application for T nonimmigrant status” and determining how many decisions there are by year (divide by precedent and non-precedent). We also exclude derivative applications.

These cases still fall into two categories -- appeals and motions to reopen and reconsider. We analyze the **appeals** category.

Appeals will say the following in the top left section of the decision: **Appeal of Vermont Service Center Decision**. Figure 3 illustrates an example of it.

Figure 9: Example of Appeal cases

In Re: 3113029

Date: JULY 30, 2020

Appeal of Vermont Service Center Decision

Form I-914, Application for T Nonimmigrant Status

We **exclude derivative applications** (applications for family members). We search within the set for cases with the following heading in the PDF document and exclude them from our general analysis. See below for this heading. The search would be for “Supplement A, Application for Qualifying Family Member of a T-1 Nonimmigrant” In the heading area below.

Figure 10: derivative applications



U.S. Citizenship
and Immigration
Services

Non-Precedent Decision of the
Administrative Appeals Office

In Re: 6361239

Date: JUNE 24, 2020

Appeal of Vermont Service Center Decision

Form I-914 – Supplement A, Application for Qualifying Family Member of a T-1 Nonimmigrant

b. How many cases were dismissed? (separated by year to demonstrate trends)

The information available at end of the decision, where it says ORDER. See below for an example. It will either say that the **motion or appeal was dismissed**.

Figure 11: dismissed cases

ORDER: The appeal is dismissed.

How many cases were granted (this can be indicated by language including **grant, sustain, or withdraw** and **remanded** back to the agency)?

We analyze by year to demonstrate trends. The motion or appeals that were granted should state at the end (final section) under OUTCOME that they were “granted” or “sustained” or “withdrawn” and/or “remanded.” Below are a few examples. The “remand” captures all of the terms because this means it is sent back to the agency.

Figure 12: granted cases

ORDER: The motion to reopen is granted. The matter is remanded for the entry of a new decision consistent with the foregoing analysis.

Figure 13: sustained cases

ORDER: The appeal is sustained.

Figure 14: withdrawn cases

ORDER: The decision of the Director is withdrawn. The matter is remanded for the entry of a new decision consistent with the foregoing analysis.

c. What grounds were cases denied? (separated by year to demonstrate trends)

In the introductory paragraph of the decision, AAO generally states the reasons for the denial. We then search in this first paragraph section for the prevalence of the following terms, and this is an indication of the grounds for denial:

- i. “victim of a severe form of trafficking in persons”
- ii. “physically present”
- iii. “complied with reasonable requests for assistance”
- iv. “extreme hardship”
- v. “inadmissible” or “inadmissibility”

We conduct this search in **denied cases** as well as **granted cases** and then compare the results in each batch. Under the **ANALYSIS** section of the decision, we also search the terms below to determine how USCIS is deciding particular cases.

- vi. “victim of a severe form of trafficking in persons”
- vii. “physically present”
- viii. “complied with reasonable requests for assistance”
- ix. “extreme hardship”
- x. “Inadmissible” or “Inadmissibility”

Other terms of interest in this section:

- “Sufficient evidence”
- “Any credible evidence”
- “Depart” and “return” or “voluntarily” or “reentry” (including conjugations of these verbs)
- “8 C.F.R. § 214.1 l(g)(l)(i)- (v)” (which is the “on account” of requirement
- “8 C.F.R. § 214.1l(g)(4)”
- “USCIS may consider when applicants escaped their traffickers, what activities they have since undertaken to deal with the consequences of having been trafficked, and their ability to leave the United States”
- “8 C.F.R. § 214.11 (g)(2)” (physical presence)
- “Preponderance”
- “Domestic violence”
- “Smuggling” (and variations -- “smuggle” / “smuggled”)
- “Escaped a severe form of trafficking in persons before an LEA was involved” or “[e]scaped a severe form of trafficking in persons before an LEA was involved”
- “labor exploitation”

Not all decisions have subsections within the ANALYSIS section, but some do.

We search for the prevalence of the following subsections within ANALYSIS:

- xi. “The Applicant’s Trafficking Claim”
- xii. “The Applicant is Not Physically Present On Account of Trafficking”
- xiii. “Extreme Hardship”
- xiv. “Reasonable Response”
- xv. “Inadmissibility”

d. How has this changed over time?

By looking at the prevalence of terms above over time, we should be able to determine trends.

e. Any patterns?

We discovered the patterns based on our Data Analysis part.