



# DC1: Exploratory Data Analysis

(Lilly Campbell, Grace Hartley, Marium Tapal, Eleni Partakki, Zampa Provenzale)

## Contents

---

<b>DC1: Exploratory Data Analysis</b>	<b>1</b>
<b>Summary: GELMZ AGENCY</b>	<b>2</b>
<b>Zampa Provenzale: Suspicious Phone Call</b>	<b>4</b>
<b>Lilly Campbell: Voting Registration Patterns EDA</b>	<b>4</b>
<b>Grace Hartley: digging into party divides</b>	<b>6</b>
<b>Marium Tapal: News Articles and Scientific Reports</b>	<b>8</b>
<b>Eleni Partakki: Maps and Scientific Reports</b>	<b>10</b>

# Summary: GELMZ AGENCY

## GELMZ AGENCY



*Four shadowy figures meet in an alleyway, the rain falls, somewhere cats are fighting and the city smells of smoke. They were an elite private detective squad of reformed cyber criminals run by the human resources department of the global shadow government which is puppet master to all other governments. Their leader was known only as JC, and it was from these mysterious initials that they first learned of the mysterious events that had taken place in the town of Alderwood...*

*They've been at it for days now. Blisters stinging the soles of their feet as they plod the sinister sleepy town of Alderwood, where sour-faced locals took none too kindly to strangers, and the streets were always covered in sin. But after tireless data analysis, door to door questioning, and breaking into the city records hall and local newspaper the results were still inconclusive.*

*Marium chews and spits tobacco "Find anything?"*

*The others shake their heads - there were more questions for all of them than answers.*

*How long had she spent pouring over old newspaper clippings and torturing the now-retired editor with a stick of butter and needle nose pliers? How many dull nights spent watching for a sign of suspicious activity outside of the local gazette, and how many keywords had she indexed against the massive text files of the town newspaper? She spit again. Had it all been for nothing?*

*"There's something there I know there is!" Says Eleni, shuffling her notes, and looming dark against the streetlamps that lit her silhouette. She had spent the week holed up in an abandoned warehouse constructing massive cork boards of maps and names and faces attached with red string. In her past life she had been a scientist - one of the best people said, but when undergraduates in her lab began growing additional appendages (and even a head) the university could no longer turn a blind eye. But this fellow, Dr Phillip Boynton and Dr Von Reiker had at one point been her closest confidants (perhaps lovers?) but time and the stress of academia had pushed them apart.*

*Zampa dismounted zir horse and looked at the rest of the Gelmz Agency. Zampa had once held an entire hotel hostage with only an oven mitt, a trowel, and a canister of gasoline, but the case of Alderwood had pushed her farther than the screams of the clowns stuck in the convention of that hotel could ever hope to.*

*What at first seemed a cut and dry case, a mysterious laboratory with the typical amount of corruption and back-door dealings had pushed each member to their psychological and spiritual limits.*

Zampa was about ready to call it "I broke into the city hall and was able to hack into their system and swipe the 2002 phone records, one was unlabeled - a call from Canada" which ze had tracked to an empty stronghold in the Audobon where a timed bomb had been labeled 'I told you so'. With quick wit and racing hands, Zampa picked apart the bomb and shredded the documents- barely escaping with zir life.

Grace too was despondent. After an extensive investigation of the Monson family's \$2.5 million payout from the town of Alderwood, the family had made her job difficult. Lawsuits ensued, and the magnitude of their fortune obscured. That, of course, was child's play compared to the blackmailing she faced from the American Angus Association. They were relentless. A single interview she conducted with Wayne Madrell resulted in being tailed for six months. She woke to dead birds on her doorstep, ominous phone messages, and even one death wish.

"I found something..." Says a voice from the entrance to the alley. They all jumped - no small thing for such an elite group of spies.

Lilly Freakin Campbell, bad boy bank robber and patagonia model turned detective, saunters down the narrow cobblestones casually swinging a sword that was lit on fire and smoking a long thin cigar.

"Lilly" the other's cry "You look so well prepared for this meeting, and sexual!"

"I know!" Says Lilly, her voice dripping with humility.

"But what did you find."

She sheathes her sword. "It's not what I found - it's what I didn't" \*gasp\*

"In 2003 there were 1,100 new registered voters to an undeclared party, all of them men - I mean what are the odds?!"

So what exactly did the GELMZ Agency uncover? Read the rest of our report to find out more and answer questions such as *WHERE* did all of those undeclared voters come from in 2003? (see graphic) and *WHY* were they all men? *WHAT* is the deal with the Monson family feedlot \$2.5 million payout from the town? Why do they have such influence despite not living in the town? (at least are not registered to vote) *HOW DEEP* does is the shadiness of the American Angus Association, and *WHY* do they have an unnamed representative convince a 14 y/o boy to join with the promise of benefitting him "someday"?

# GELMZ Analysis

## Zampa Provenzale: Suspicious Phone Call

### Suspicious phone call to Canada January 26th 2002

In January 2002 only 1 call at City Hall was made outside of Washington; to an area code matching Toronto, Canada. This was also the only call made during a weekend and the only call without a name/business attached. This call was made on January 26th 2002 but we have no additional information. I noticed that the scientist “relocated,” we don’t know from where. The call was made *before* the economic collapse. However the phone log only has data for January 2002. We cannot compare this to trends in other months, see if the same number was called at other times, or if there have been any other unidentified weekend calls. This could easily be a coincidence, but might not be.

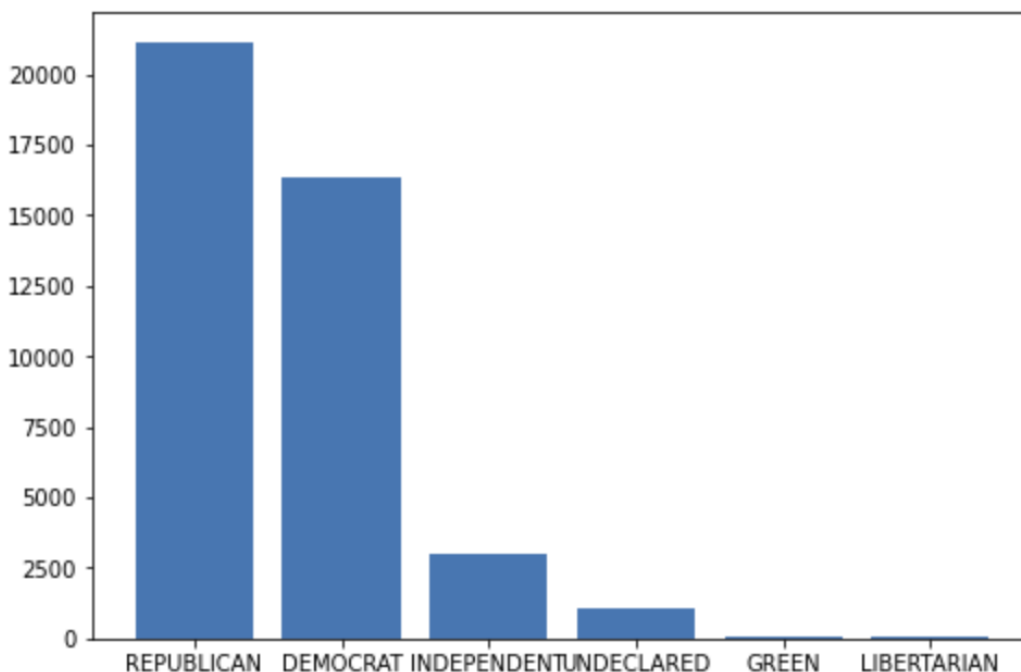
UPDATE there is one article that mentions “Toronto” (thanks Marium!!) In February 2003 the city hosted an ensemble from Toronto. The mystery call was probably to schedule this showing. But why was it on the weekend? Why was there no name attached? WHY do we only have one month worth of phone logs?????

## Lilly Campbell: Voting Registration Patterns EDA

### Process/Guiding Questions

Given the voter registration data dump (including years 1935 - 2004, date of birth, sex, ssn, and street of residence) can we parse any interesting findings about voting patterns in this town?

Overall here is the aggregate party distribution



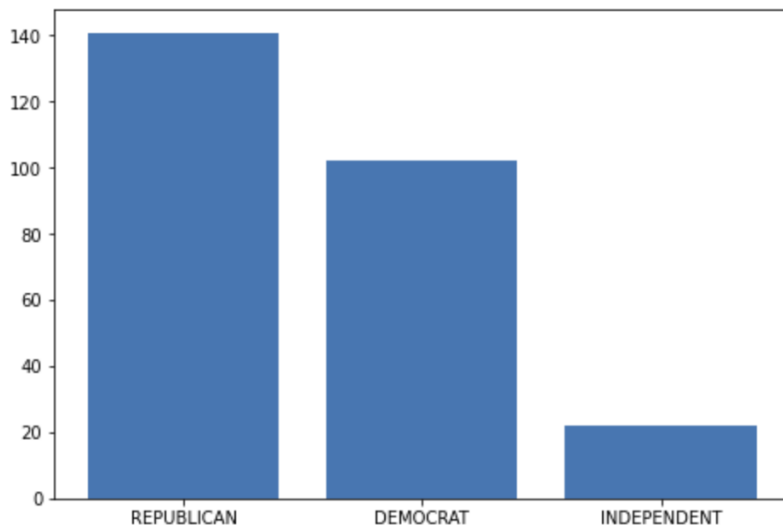
With the following count for each group:

PARTY	
DEMOCRAT	16374
GREEN	40
INDEPENDENT	2996
LIBERTARIAN	39
REPUBLICAN	21170
UNDECLARED	1100

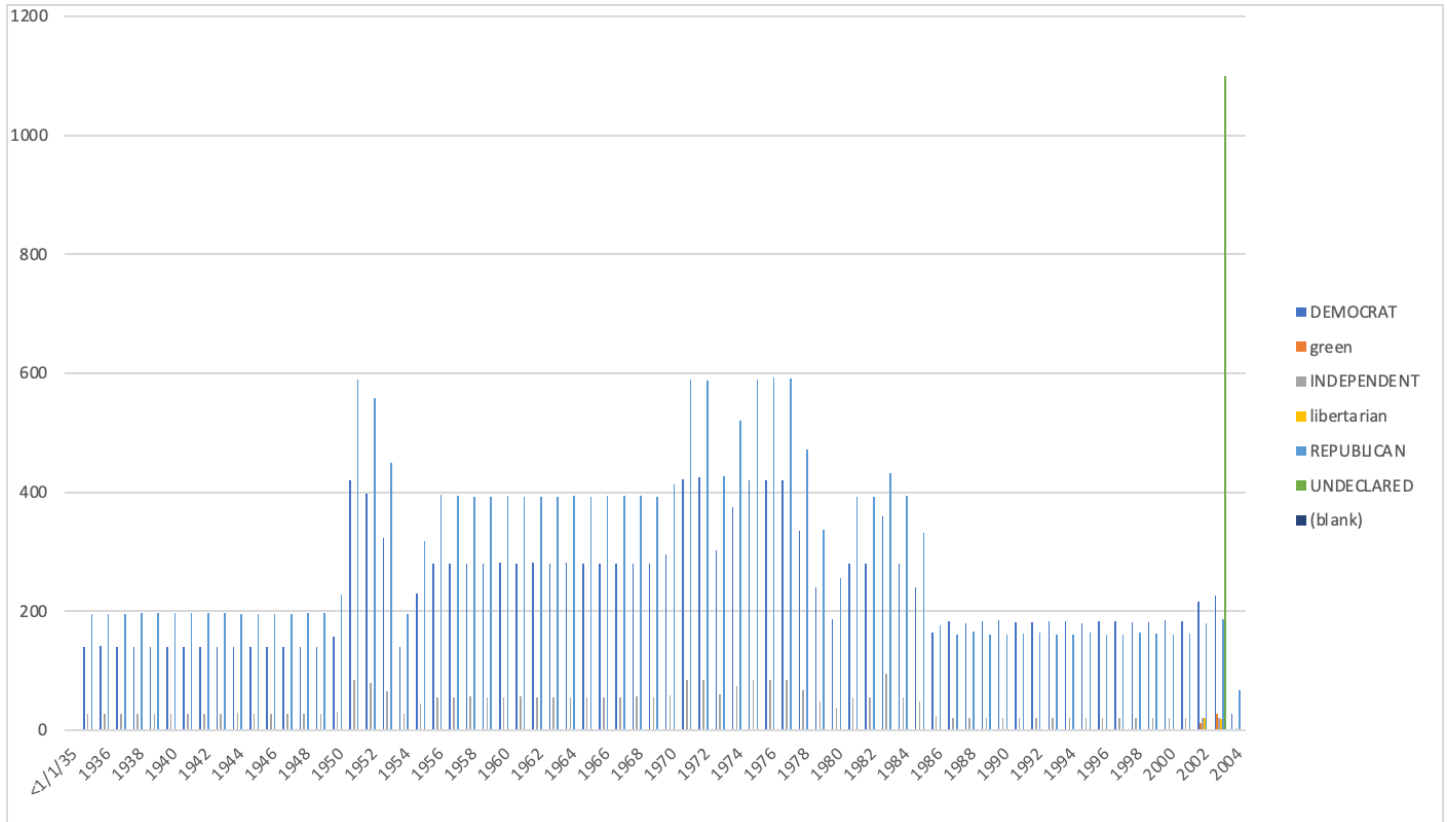
I then grouped by street names and parties to see if there were any standouts

I wanted to look at Oak st in particular since this is near the laboratories

PARTY	
DEMOCRAT	102
INDEPENDENT	22
REPUBLICAN	141



Doesn't seem too different...



I decided to play around with the party activity in excel and it seems like there is a HUGE spike in undeclared voting registration around 2003 so I'm going to continue to examine this..

## Grace Hartley: digging into party divides

	PARTY <chr>	num_born_in_ald <int>	num_transplants <int>
1	DEMOCRAT	7846	8448
2	REPUBLICAN	10472	10698

2 rows

Looking along party lines, it doesn't seem like place of birth (being born in Alderwood vs. not being born in Alderwood) makes a huge difference. It doesn't seem like there is a political divide between people who are from Alderwood or not.

Family ties:

LAST NAME <chr>	N <int>
JOHNSON	337
SMITH	332
ANDERSON	280
JONES	177
MILLER	177
BROWN	157
THOMAS	137
LARSON	130
NELSON	130
WILLIAMS	130

1–10 of 3,365 rows

Previous 1 2 3 4 5 6 ... 100 Next

These are the most common last names in Alderwood, based on the Voter Registry. We may be able to trace the biggest families with the largest influence from here, though we can't necessarily assume that people with the same last name are related (particularly with such common last names).

With this in mind, I decided to disregard the top seven most common last names (a completely subjective and not systematic process). I searched the news articles for mentions of the names "Larson" and "Nelson", and found some promising leads:

- Larson gallery - Yakima Valley Community College
- Laura Nelson - selected as 1 of 2 co-alternate state dairy ambassadors (along with Kristen Irving. McKenzie Klein was chosen as the captain to represent the dairy industry). The position has a 10-month paid internship, \$4,500 in scholarship money, and training in communications.
  - Candidates must have dairy knowledge either from family business or other independent work experience - **are the larson's in the dairy industry?**
- Larson family and Edes family are connected by the marriage of Carolyn Edes to Ernie Larson (via Margaret May Bolin Edes obituary)
- Alene Nelson of Texas - sister of Hazel M. Hannaman of Grandview(via the latter's obit)
- Mark and Renee Nelson (Grandview) join the **American Angus Association** in Dec 2003
  - AAA is largest beef cattle registry association in the world
  - **Ellie Olmsen** is the local point person
- **Nelson Irrigation Corp.** based in Walla Walla - known worldwide for "innovative water-conserving lawn and agriculture sprinkler systems"
  - Connected to some conservation news + Walla Walla Watershed Alliance in October 2004
- **Monson family - sold its 150 acre feedlot to city in Oct 2004**
  - Seen as opportunity for economic growth in Alderwood and improvement in quality of life
  - \$2.5 million payout from an internal loan
  - **Kenny Nelson** of DK Bain Real Estate

- Excited that people might want to move to Alderwood instead of Grandview of Zillah if no more cow smell + dust
  - BUT they still will keep their 20 acre adjoining lot
  - Monson's have to clean up their site
  - [There are no Monsons in the voter registry](#)
- Wayne Madrell (14 y/o)
  - Resident of outlook who used money he earned from selling livestock via FFA and 4-H to start his own herd of 4, wants to grow more
  - Madrell family part of cattle industry
  - Member of **American Angus Association**
    - **Because the person who sold him the cattle convinced him it would “help him someday”**
    - AAA -> Better breeding services -> higher market
  - This story is by **Ellie Olmsen**, see the Nelson AAA point
- Alderwood not raising property taxes - 11/2004
  - People think its to fund buying the Monson feedlot
  - **Tom Seeger**
  - **Jim Torch**

## Marium Tapal: News Articles and Scientific Reports

### Summary:

- This story is playing out in two main locations: Alderwood, Washington and Switzerland. We suspect that there is a connection between the Swiss Von Ryker Institute (the institute researching the Mad Cow Disease) and the “Boynton Labs” in Alderwood. A secondary location that might have links to the scenario is Argentina, where the first case of the Mad Cow Disease was found.
- The sentiment analysis bar plot (see graphic below) tells that the general sentiment of the news in Alderwood is positive. In addition, there is also a lot of anticipation of what is happening in the town (perhaps relating to the new facility opening), and trust in the city officials.

### My work process:

- I used base R and the following packages for my analysis: tidyverse, stringi, readtext, tm, SentimentAnalysis, syuzhet, wordcloud.
- I started out with reading the Scientific Reports and connected them with the News Articles to find more information.
- I also performed a simple text and sentiment analysis following this [tutorial](#). I thought this might be useless but it proved to be useful.
- I helped my group members find relevant articles to figure out the details for their work.

### What I found:

- **Locations in the story:**

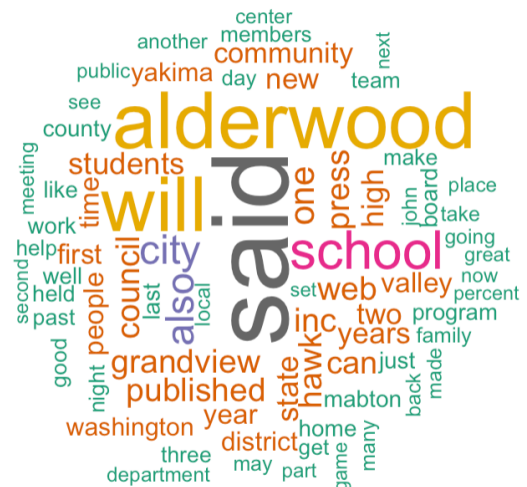


- Alderwood, Washington
- Switzerland:
  - The scientific articles mention things about the Swiss Von Ryker Institute AG and the mad cow disease (a neurological disease). In the scientific article 4, it is mentioned that the institute closed. This is based upon the allegations of bioethics violations. This institute researches protein structures of neurological diseases.
  - It feels like there is a connection between the new private laboratory in Alderwood facility specializing in agri-business and the mad cow disease that the Swiss institute was looking at)
- Argentina: Quick search for “labs” in the news articles found a certain Boynton Labs which looks like has links to Argentina (Article: 1101163018612.txt)
  - Boynton Labs was chosen for testing the Mad Cow disease by the US Department of Agriculture Animal and Plant Health Inspection Service. The first case of the Mad Cow disease was discovered. Beef exports from Argentina are being banned in countries including The Philippines, Japan, South Korea.

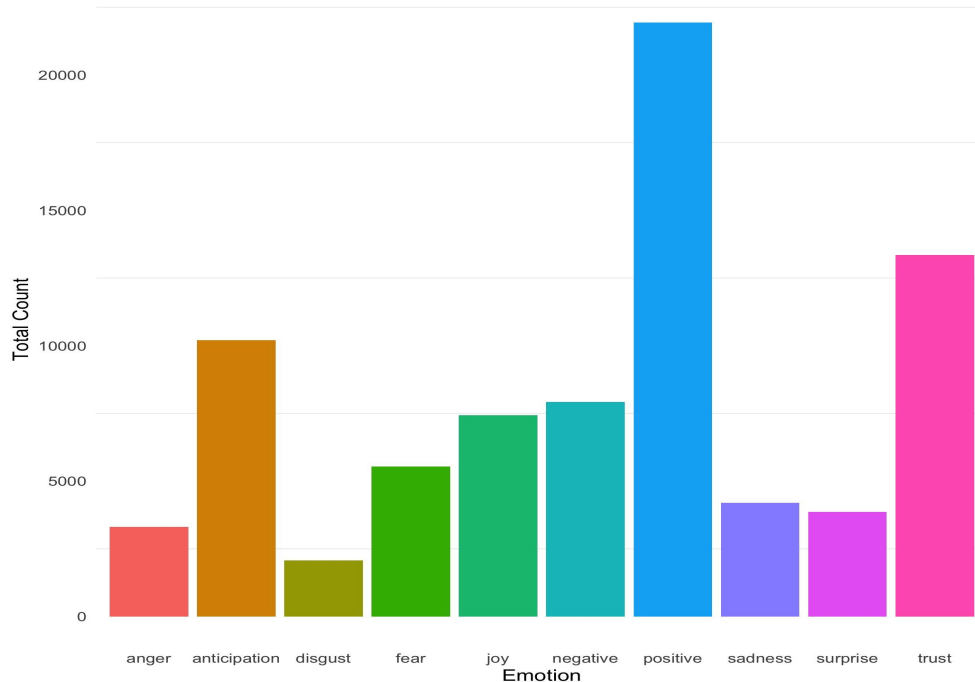
● **Simple Text Analysis of the News Articles:**

- The word cloud below shows the top words being mentioned in the news articles
- Some interesting words we could look into are mabton, hawk, yakima, john (what/who are these?) School is repeated often - what happens at the school and which school?
- Manual search and looking into indicates:

- **Mabton:** Highway? (1101162470201.txt). Also mentions Mabton High School (1101162489920.txt), and it's grid team. I can't find Mabton on the city and state map. Maybe it is in a neighboring city - rival to Alderwood HS.
- Alderwood High **School**.
- Roberson and her sister who lives in **Yakima** (1101162469467.txt) (a place in Alderwood)
- Hawk: can ignore - that's the name of the press company publishing these articles
- John: also can ignore - that's the name of the journalist writing these articles



Sentiment of Alderwood News Articles



- The bar plot shows the general sentiment of the news articles. Lots of positivity but also anticipation of what is happening in the town, and perhaps, trust in the city of Alderwood officials.

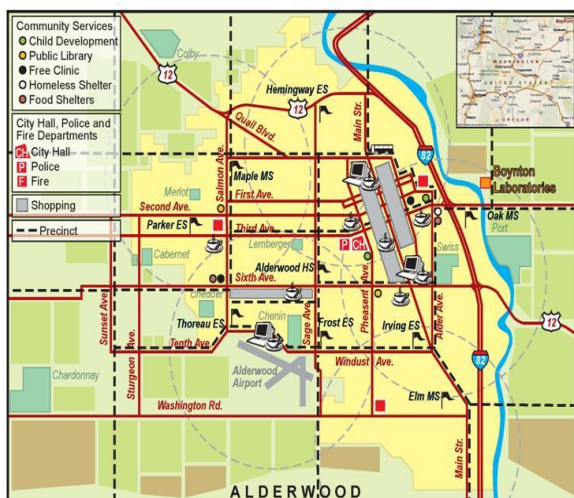
## Eleni Partakki: Maps and Scientific Reports

As the rest of the GELMZ agency took on the data, I focused on the two jpeg file maps and the scientific reports.

### To begin with the maps:

Map of Alderwood, Washington, city view: Some key features of the map are the precincts.

Map of Alderwood, Washington, city view



We can notice some distinct precincts through the square and circular dotted lines on the map. We can see that Boynton Laboratories are on the other side of the water body separating the main city of Alderwood. Some may claim that this is simply a result of protecting the city by potential experiments, but I am not convinced. In addition, the lab is separated from the center of the city by a highway, making it easier to hide whatever it is that they may be working on. It is vital to note that Alderwood is in a very strategic location, in the center between three main cities, Portland, Seattle, and Spokane. Could there be a link? Are

there any exports in relation to these cities? If we had more time, I would have wanted to find out.

For anyone to leave the lab, they would have to cross the body of water by boat, or by taking the road to Route 82 or by driving through roads to Route 12, then meet an intersection at Route 82, and be able to exit Alderwood without even having to interact with the city. I wish we had more geospatial data. If someone from the lab had to interact with the city, they would be met with what we assumed to be cafes and shopping, but two of the closest community services involved the homeless and food shelters. The lab is relatively tucked away and able to hide.



## The link between maps and information mentioned in the scientific reports:

Dr. Serraet suggested that: "I would advise the managers of cattle abattoirs to sell the meat only to lower-class butcher shops." This activated a question for me. Were they running experiments in the lab and then testing them on the lower-class districts that were easily accessible to them?

This sparked my interest and I collected names that I thought would be important to look further into from the reports.

### The names that stood out were:

- Dr. Pierluigi Gambetti
- Dr. Stanley B. Prusiner
- Viadero
- Dr. Serraet
- Von Ryker
- Delwin Sanderson

Some names and keywords like "funds" brought up way too many results when I connected with Marium.

If we had more time, I would have looked into what transactions there are between Boynton Laboratories and the city.

**An issue that continued to pop up in the articles is how the disease was not initially an issue in the US:**

"In the United States, about one person per million develops and dies of the sporadic form each year, meaning that there are usually about 300 cases in the country. (For unknown reasons, Italy has nearly two cases per million; Switzerland has about three.)

Only one person in the United States has had variant Creutzfeldt-Jakob, the human form of mad cow disease, which was first described in the 1990's and has killed about 150 Europeans; she grew up in Britain and was probably infected there. "

As a result, this makes me believe that someone injected this issue into the Alderwood community. There question is, WHY? Dr. Serraet suggested that solving the problem would be easy by picking who to feed infected meat. But is it that simple? Or is this a test?

**Lastly, we are informed that:**

"Six months before Viadero left his job as inspector general of the U.S. Department of Agriculture, a shipment of 650,000 pounds of foreign pork arrived on the East Coast. Government inspectors apparently approved the shipment and sent the meat on its way to America's heartland.

At its destination in Iowa, a meat inspector noticed a serious problem. The pork came from a country that was already under a U.S. embargo because of the threat of foot-and-mouth disease. Viadero won't say which country was the source. "The mere fact that it got in here," he said. "That's the big issue."

Viadero is suspicious. He knows the source country, but did not want to inform the world where this infected foreign pork came from. What does Viadero have to do with imports of ill meat? And can we somehow link him to Alderwood? And could Viadero have something to do with Von Ryker and the investigation for ethical misconduct?

Von Ryker was proved to have infected over 100 bovine, with questionable methods. The Von Ryker Institute was closed, but could that have been dust in our eyes?

**Sanderson used to work with Von Ryker, and suddenly he is in Alderwood?**

We had traced him with Marium in the News Articles, commenting on the meat at a new restaurant "Chez Becky". *"Is she Belgian? She uses a lot of Worcestershire. But I like Worcestershire. I will be back."* notes Delwin Sanderson, off to catch a flight to South America for another conference. *You won't find this kind of food in Argentina, Delwin!"*

Why is Sanderson in Argentina? Is he hiding? And what kind of conference could there be in Alderwood after the chaos?

Something tells me that Von Ryker and Sanderson might be in Alderwood for messy business, but for now, I must research more. GELMZ is still on mission.