# Training codebook

## Claire Ewing

## 11/13/2020

## Codebook

To assess whether the violator is an individual or company (`indiv_company`):

Variable type: nominal (levels: individual, company, municipality, other, multiple)

1. Look at the name of the offender. If it's the name of anything except for an individual (first name, last name or last name, first name, depending on the province), it's an industrial offense. If it's an individual's name, safe bet that it's an individual offender, but double-check the extended description to see if they are taking the fall for their company. If they are taking the fall for their company, we'll call it a company offense.

- Examples:

    – Easy: "Name of offender: Norbert Kossler; Extended summary: Burning debris so as to reduce visibility near airports or highways" -BC data, Air, line 5 (Norbert Kossler). The name of offender is in first name last name format and there is nothing in the extended summary to suggest Norbert is taking the fall for a company. It's a straightforward individual offense.

    – Trickier: "Name of offender: Baich, Robert; Extended summary: RVB Management Ltd. and Robert Baich conducted an activity, the unauthorized burning of prohibited debris on November 26, 2010." -Alberta data, Enforcement Quarterly Reports, line 894-895 (Baich, Robert). So if you were only assessing whether the violator is an individual or company based on the name of the offender, you would probably assign this as an individual, since it's in last name, first name format. But when you look at the extended description, you see that it was RVB Management as well. A quick Google (and context clues and common sense) will tell you that this is Baich's company, and he's taking the fall for this action. Accordingly, we should actually code it as a company offense. (http://www.salespider.com/bp-53194906/robert-baich)

To assess the time span of an offense's **pollution** (`pollution_time_span`):

Variable type: dbl 1. This is for offenses that occurred over a given span of time. For many of these occurrences, I have already "dirtied" the date of offense column by putting the time span in there when I was creating the dataset. In order to clean the date of offense column, that time span can be made into a quantitative variable in this new column, calculated by the number of days an offense continued. These date ranges might be regulatory or pollution related, so check the description to ascertain which they refer to. Pretty sure the function `ddays` in the `lubridate::` package can do this, but I haven't used it before.

2. However, I have not done this to all observations where a series of dates are involved, so you'll need to double-check the extended summary for all observations to see if information about an extended time series is included.

3. You'll notice that there are relatively few occurrences where the date range is clearly or cleanly ascertained, either by me already or as you'll find in the extended summary. Put NAs where there is no time span to be found.

- Examples:

– Easy: "Offense date: 9/15/14 - 7/25/15; Extended Summary: Discharged dust at a property that is not an approved waste disposal site" -Ontario data, Bulletins, line 8 (Siefker Inc.; Dale Siefker). There are 314 days between these 2 dates, so that's the time span of the offense (as estimated by the court). 314 days would be this column's input. Additional exercise, would this be an individual or company violator? (Company - read the bulletin linked in the doc if you're curious).

– Trickier: "Date of enforcement: 12/18/07; Extended summary: The Company operates the Fort McMurray Oil Sands Processing plant pursuant to an Approval. The Approval requires the Company to conduct, operate and maintain, and report on ambient air monitoring. During the last eleven months there have been many instances where the measured values of H2S in the ambient air at stationary air monitoring stations downwind from the plant have exceeded the Alberta Ambient Air Quality Objectives for H2S. It had several incidents in which there have been releases of substances that caused or had the potential to cause an adverse effect. The ground level concentration (GLC) Exceedances and Emissions Management Program that began in November 2006 has not resulted in improved environmental outcomes. Alberta Environment ["AENV"] is of the opinion that the Company is contributing to regional GLC exceedances measured as H2S which may cause adverse effects. It shall continue the work proposed in its GLC and Emissions Management Program; develop, submit a Monitoring Plan; develop and submit a Modeling Plan; develop and submit a Diluent Management Plan; immediately restrict diluent losses to the tailings ponds to rates of 1400 bbls/day; and submit written progress reports on a monthly basis." -Alberta data, Enforcement Quarterly Reports, line 563-579 (Suncor Energy Inc.). Operative phrase: "During the last 11 months". So you can approximate the number of days in the 11 months since Nov 2006 (~335 days).

To assess the time span of an offense's **regulatory behavior** (`regulatory_time_span`):

Variable type: dbl

(See notes 1-3 for `pollution_time_span`)

- Examples:

  – Easy: "The Company operates the Hansman Lake Sour Gas Plant pursuant to an Approval. It contravened its licence by failing to take air monitoring samples for the months of May and June 2008 as required in its approval." - Alberta data, Enforcement Quarterly Reports, line 610-612 (Husky Oil Operations Limited). The regulatory behavior offense spans 2 months, thus input 60 days as the value in this column.

  – Trickier: "Failed to characterize the sources of odor emissions in the year following the commissioning of the treatment of digestate, to produce and transmit within 30 days after the end of the year of characterization an odor dispersion study" -Quebec data, Both Registers, line 366 (Ville de Saint Hyacinthe). Not necessarily trickier, just hard to read. They were supposed to submit the report wihtin 30 days of end of year, so it'd be "30".

To assess whether an offense's **pollution** is acute (accidental, more or less a one time deal, brief blip) versus chronic (ongoing, seems they were doing this regularly or continuously and finally got caught on x date): `pollution_acute_chronic`

Variable type: ordinal (levels: very short <1 week, short >1 week <2 months, medium >2 months <1 year, long >1 year <3 years, very long >3 years, unknown, and multiple for multiple violations)

1. For where there is no time span described, we want to get some sense of the time scale from the extended summary. We're loosely calling this "acute versus chronic", but as described above, the levels are more specific.

2. There will be lots of NAs in this column as well where we are unable to infer durations from the extended summary. It will be particularly difficult with Quebec, where the extended summaries are brief, but should be easier in places like Alberta, where they contain more detail.

3. It will be tricky/not as intuitive as the other categories, so let me know if this is an impossible task and we can figure something out.

4. When dealing with multiple violations within a single observation, assess which categories are multiple. For example, if there is a single pollution violation and multiple regulatory violations, write the more exact level for the pollution violation and write multiple for the regulatory violations.

- Examples:

  – Easy: "The Company burned prohibited debris, which included a gutter, a metal barrel, and plastic or aluminum siding, without the proper authorization." -Alberta data, Enforcement Quarterly Reports, line 507-508 (GDW Contracting Ltd.). I would say it's a safe bet that most individual-level and most company level burns, especially those with small fines, are "very short". Still read the extended summaries and see if there are any burns that would lead you to believe that they burned for longer than a week, but I'd be willing to bet most burns will be "very short".

  – Meh: "Extended summary: Failed to install devices to detect and record leaks from dust collectors" -Quebec data, Both Registers, line 114 (Lafarge Canada inc.). I would say this is a "medium", as with most of the Equipment "failure to install" type questions.

To assess whether an offense's **regulatory behavior** is acute (accidental, more or less a one time deal, brief blip) versus chronic (ongoing, seems they were doing this regularly or continuously and finally got caught on x date): `regulatory_acute_chronic`

Variable type: ordinal (levels: very short <1 week, short >1 week <2 months, medium >2 months <1 year, long >1 year <3 years, very long >3 years, unknown, multiple for multiple violations)

(See notes 1-4 of `pollution_acute_chronic`)

- Examples:

  – Easy: "The Company operates the Taber thermal electric power plant pursuant to an Approval. It contravened its Approval by submitting the 2005 Annual Air Emissions Report late; and failing to immediately report by telephone any contravention of the terms and conditions of the Approval." - Alberta data, Enforcement Quarterly Reports, line 490-493 (Maxim Power Corp.). This would be "multiple" regulatory_acute_chronic, because it is 2 occasions of Notification violations occurring on separate timelines, and "NA" for pollution_acute_chronic, because we have no data for the latter.

  – Trickier: "Failed to register the use of and calculate the emissions of VOCs while painting" -Quebec data, Both Registers, line 85 (Royal Group Inc.). I would say this is "short" regulatory_acute_chronic - I imagine that's a fair timeline to register emissions and for a big enough painting project that you have to register VOC's, I imagine it's going to take longer than a week. This is basically a judgement call. Exercise "rationalizing" such calls as I have done here.

  *Sythensis example* - both regulatory and pollution acute_chronic example:

  – Easy: "Failed to notify ministry of a discharge of natural gas into the environment" -Ontario data, Bulletins, line 33 (Union Gas ltd et al.). `regulatory_acute_chronic` would be "very short", because it is especially clear when you read the bulletin, it's clear that the discharge was accidental and discovered quickly (happened when they were digging), and thus could have been reported in a short time scale (i.e. within the week). Similarly, the nature of the accident was very quick, so `pollution_acute_chronic` would also be "very short".

  – Trickier: "Offense date: 9/20/12; Extended summary: The Company operates the Drayton Valley Thermal Electric Power Generating plant pursuant to an Approval. It contravened its Approval when it released particulate matter at a dry concentration into the atmosphere in excess of the limit and exceeded the mass emission limit; late submission of the August 2010, December 2010, January 2011 and February 2011 monthly air emissions source monitoring reports; release of carbon monoxide at a dry concentration into the atmosphere in excess of the limit; and failure to immediately report a contravention of the Approval. Paid 15 October 2012." -Alberta data, Enforcement Quarterly Reports, line 863-870 (Valley Power Corp.).

To assess keyword (AKA the briefest possible phrase to describe and group together the observations by their extended summaries) (`keyword`):

Variable type: nominal

1. Especially to distinguish between Excess emissions and Standards in Quebec, but also to gain more details generally, look at the articles and regulations. You can just google the letter/number combination and "quebec" and you can poke around on the http://legisquebec.gouv.qc.ca/ website to get more details. This will help you distinguish whether an extended summary that says "excess particulate emissions", for example, was in excess of the company's permit (making it "Excess emissions") or in excess of the province's standard (making it "Standards").

Levels:

- "Burn" - when an individual or company burns something (unauthorized)

  *Examples*:

  – Easy: Can be simply "burn" in the description, or "The Colony conducted an unauthorized activity, the burning of prohibited debris including plastics, household garbage, tires, treated wood, steel and other items on September 28, 2017." -Alberta data, Enforcement Quarterly Reports, lines 957-959 (Hutterite Brethren of Wilson Colony). Self-explanatory.

  – Trickier: "Description: Introduce Business Waste; Extended Summary: Community Environmental Justice Forum: A Forum was held in 100 Mile House in response to the October, 2009 burning of prohibited material in a beehive burner by Ainsworth Lumber Co. Ltd. During the Forum, which brought together representatives from the company and the community, appropriate restitution for the incident was identified. The company was required to pay $20,000 to establish a local bear aware program and $4500 to fund the operation of an air monitoring station within the town site for three years. Additional restitution included issuing a public apology in the local newspaper, posting signage, and conducting instructional staff meetings." -BC data, Air, line 10 (Ainsworth Lumber co. Ltd.). Note how the short description describes it as "introduce business waste", but the extended summary shows that it is in fact a "Burn". This goes to show the importance of reading the summary! Also, note how BC's data only has years and quarters listed as a time frame, but when you read the extended summary, you can extract October 2009 as the offense date and add that into a new column in that dataset.

- "Equipment" - when an individual or company (usually company) fails to install air pollution (abatement) equipment, doesn't use it properly, fails to repair it, doesn't upgrade it or maintain it in accordance with standards, uses it without permission, etc.

  *Examples*:

  – Easy: "Failed to install or properly operate ammonia scrubber" -Quebec data, line 37 (Anco chemicals inc.). A scrubber is a piece of air pollution abatement equipment.

  – Trickier: "Pollution Abatement Order issued to cease the operation of a wood boiler after numerous public complaints were received by the Ministry and the local fire department regarding excessive smoke and noxious odours. The proximity to a school yard and residential area made the operation of the improperly functioning boiler not possible without causing pollution and harming human health and the environment." -BC data, Pollution, line 6 (Diana and Richard Gravells). This isn't necessarily trickier, just more elaborate (couldn't find any really tricky ones). Might be a bit confusing with the smoke/odours bit, but clearly it is the wood boiler equipment that is causing the problem, as it is not functioning properly.

- "Monitoring" - when a company (don't think I've seen any individual level cases, but still check) fails to document, maintain a register/record of, monitor, etc. their air pollution

  *Examples:*

4

– Easy: "Failed to sample ambient air for atmospheric emissions" -Quebec data, line 51 (Vitaliterre inc.). It's probably part of their permit to sample and report their emissions, so failure to do so resulted in this penalty.

- "Notification" - when an individual or company fails to notify the relevant government body (usually Ministry) that they have emitted x amount of air pollution, when they have exceeded their emissions permits, when they are going to be late with their pollution reports, etc.

  *Examples:*

  – Trickier: "Information Order issued to submit all emission testing done on this faculty since its inception, as required under the Asphalt Plant Regulation (APR). The Ministry currently has no documentation of any stack sampling performed at this site, which is a contravention of the APR. The 2006 paving season resulted in a large volume of complaints about the asphalt plant operations and air quality concerns." -BC data, Emission, line 8 (RD Moyen Holdings Ltd.). Again, not terribly tricky, just more elaborate. It might look like it's Monitoring, because of the lack of emission testings, but the actual problem is that the Ministry has no record of it (they have not been notified of it). Thus it is a Notification coding.

- "Standards" - when an individual or company (usually company) fails to comply with a given air pollution standard

  *Examples:*

  – Easy: "The company being the operator of an incineration plant on behalf of Quebec City, emitted combustion gases into the atmosphere containing more than carbon monoxide than the prescribed standard" -Quebec data, Both Registers, line 349 (Tiru (Canada) Inc.). Most of the Standards observations are straightforward-ish like this.

  – Trickier: "Emitted excess dioxins and furans" -Quebec data, Both Registers, line 140 (City of Quebec). This looks like a classic "excess emissions", it says it in the description, but upon examination of the articles of the law, it is actually a "Standards" keyword. "Subject to section 133, the incineration facilities governed by this Chapter must not emit combustion gases into the atmosphere that contain more than 0.08 ng/m3 of polychlorinated dibenzofurans and polychlorinated dibenzo [b, e] [1,4] dioxins." (this isn't all of it, but it's the relevant part of the law). Note that it is a "universal" ceiling/standard of 0.08ng/m3. Therefore, it would actually be a "Standards" coding.

- "Order" - when an order is issued but either the noncompliance is not clear or has not yet occurred (thus other keywords are not appropriate)

  *Examples:*

  – Trickier: "Information Order issued for non-compliance with the Asphalt Plant Regulation. The company was ordered to: (1) conduct specific stack monitoring and evaluate results to indicate whether emission criteria are met; (2) notify the Ministry at least 24 hours prior to the start-up for stack sampling purposes; (3) submit all stack sample results to the Ministry within 14 days of testing; and, (4) notify the Ministry and cease production should any sample fail to meet emission requirements." -BC data, Emissions, line 7 (Pittman Asphalt and YCS Holdings Ltd.). The standard here is the Asphalt Plant Regulation, which the company is in non-compliance with (don't get bogged down in what comes next in the description, that's what they're ordered to do, not what the company already did).

- "Excess_emissions" - when a company emits more of a given pollutant than it is permitted to, or emits any pollutant if it holds no permits at all

  *Examples:*

  – Easy: "The Company operates the Scotford cogeneration power plant, air separation unit, and carbon dioxide purification and liquefaction unit pursuant to an Approval. It contravened its Approval by exceeding the maximum hourly limit for nitrogen oxides from the Gas Turbine/HRSG Exhaust

Stack. Paid 1 June 2004." -Alberta data, Enforcement Quarterly Reports, line 144-149 (Air Liquide Canada Inc.) The `keyword` is Excess_emissions, because it exceeded its maximum hourly limit for nitrogen oxides, and I could find no evidence in the law articles of it being a universal standard. Article 227e which corresponds with this observation suggests that it is specific to the company's permit.

- "Emission" - when an individual or company emits a substance. This is sort of a catch-all, use sparingly, for when extended summaries just say "emitted dust", for example, and we don't know whether that was excess (unpermitted) or above a standard or not. Double-check the acts and regulations for observations when looking at these

  *Examples:*

  – Easy: "Discharged sulphur into the atmosphere" -Ontario data, Bulletins, line 49 (Safety Kleen Canada Inc.). I looked at the bulletin again and poked around in Ontario's environmental laws, and couldn't find anything to help indicate whether this was a permit violation or a standards violation. Accordingly, it can go in the catch-all "Emission" category.

-"Multiple" - when a single observation/enforcement action captures multiple infractions. When you see this, simply put "multiple" as the keyword (don't put the number of infractions or details about these infractions)

*Examples:*

–Easy: "The Company operates the Oil Sands Processing Plant and Mine pursuant to an Approval. It contravened its license by releasing vapours containing nitrogen, hydrocarbons and H2S through ineffective operation of the VRU hydrogen sulphide treatment unit for the periods of August 26th through August 31st, 2007; and failure to immediately report this contravention of the Approval. Notice of Appeal received 13 November 2009 by the Environmental Appeal Board.Amended Notice of Administrative Penalty issued 5 February 2010 reducing the total penalty by $2500.00 for the reporting contravention. Paid in full ($5000.00) 8 February 2010. Discontinuance of Proceedings received by the EnvironmentalAppeal Board 10 February 2010. File Closed." -Alberta data, Enforcement Quarterly Reports, line 691-700 (Suncor Energy inc.). Emission, Equipment, and Notification keywords would all apply here across the three infractions, so code as "multiple".

–Trickier: "Failed to comply with the standard for particulate emissions; Emitted particulates such that it was visible more than 2 meters from the point of emission" -Quebec data, Both Registers, line 184 (Carrieres Quebec inc.). Note that there are two violations occurring, two different Standard violations (look at the laws).

To assess `number_of_violations`:

Variable type: dbl

By definition, this will always have a value (will never be an NA) and will always be >= 1. For the observations where you coded a keyword other than "Multiple", the number of violations should simply be "1". For the observations where you coded the keyword as "Multiple", the number of violations should be a count of violations as deduced from the extended summary. To use the same examples from the "Multiple" keyword section:

*Examples:*

–Easy: "The Company operates the Oil Sands Processing Plant and Mine pursuant to an Approval. It contravened its license by releasing vapours containing nitrogen, hydrocarbons and H2S through ineffective operation of the VRU hydrogen sulphide treatment unit for the periods of August 26th through August 31st, 2007; and failure to immediately report this contravention of the Approval. Notice of Appeal received 13 November 2009 by the Environmental Appeal Board.Amended Notice of Administrative Penalty issued 5 February 2010 reducing the total penalty by $2500.00 for the reporting contravention. Paid in full ($5000.00) 8 February 2010. Discontinuance of Proceedings received by the EnvironmentalAppeal Board 10 February 2010. File Closed." -Alberta data, Enforcement Quarterly Reports, line 691-700 (Suncor Energy inc.). This would be 3 violations, so put "3" in this column.

–Trickier: "Failed to comply with the standard for particulate emissions; Emitted particulates such that it was visible more than 2 meters from the point of emission" -Quebec data, Both Registers, line 184 (Carrieres Quebec inc.). This would be 2 violations, so put "2" in this column".

To assess pollutant type (`pollutant_type`): what type of substance is of concern in the observation (what was emitted/failed to report/notify/standardized/excessively emitted)?

Variable type: nominal (levels: SO2/sulphuric dioxide, mercury, particulates, dust, benzene, odour, ammonia, VOC (volatile organic compounds), etc.)

1. This is especially relevant for the "Standard", "Excess_emissions", and "Emission" categories. Most of the time, an extended summary will say "x company/individual emitted (too much) y substance". That "y substance" is what we're putting in this category.

2. I'm not going to list out all of the possible levels here, because that would basically involve me going through all of the observations and just doing this task myself! I know that all of those listed above make at least an appearance or two in the provinces, but our training will enable you to recognize the context around other chemical etc. names and assign them to this category accordingly. (Some observations randomly get really specific with the chemical names!)

- Examples:

   – Easy: "Emitted dust" -Quebec data, Both Registers, line 132 (Minerals Mart inc.). The pollutant type would be "dust". A lot of the observations in the Quebec dataset are that simple! – Trickier: "Discharge of benzo(a)pyrene, a contaminant that exceeded established standards, for violating 3 ministry approval conditions, and for alteration of equipment without ministry approval" -Ontario data, Bulletins, line 4 (Ingram Asphalt inc.). The substance would be "benzo(a)pyrene".

To assess type of penalty (`penalty_type`): what type of penalty was issued? A fine, a warning letter, a pollution abatement order? Etc. Variable type: nominal (levels: fine, warning, letter, pollution_abatement_order, enforcement_order, pollution_prevention_order, open_court_proceeding, restorative_justice, information_order, court_conviction, administrative_penalty, long-form_information, any new ones you may come across)

- Examples:

   – Easy: "Fine: $65,000.00; Victim Fine Surcharge:$16,250.00; Extended summary: Discharged chlorine gas into the atmosphere and failed to notify the ministry" -Ontario data, Bulletins, line 5 (Priestly Demolition inc.). This is a court conviction The description does not indicate any additional penalties.