

# USD Coding Procedures

Latest revisions by Karim Bahgat  
7<sup>th</sup> June 2017

This document describes the established steps and procedures that coders should follow when coding events for the Urban Social Disorder (USD) dataset, with and without the USD Coder software.

## Contents

Before you begin .....	2
What you need.....	2
Searching for articles.....	3
Going through the reports .....	4
Finishing up.....	6

## Before you begin

1. **Determine which city and time period** you will be coding. The selection of cities to be coded should be in accordance with the prioritized cities in “Cities.xlsx” and coding history in “Coders.xlsx”.
  - a. The main selection criteria is that we code the capital city of each country, limited to capital cities of above 100,000 inhabitants, including any city that was once a capital in the period since the start of the dataset (1960).
  - b. In addition, we include the country’s largest city if its population is more than twice as big as the capital city.
  - c. For some major countries with multiple mega-cities we may code additional, select major cities.
2. If coding a new city that has not been coded before, **register the new city in the “City Listy.xlsx” file** (the “Cities” sheet). This includes assigning:
  - a. A unique city ID.
  - b. The country in which the city is currently located.
  - c. The continent/region of the city.
  - d. The approximate latitude and longitude coordinates of the city’s centroid.<sup>1</sup>
3. **Get the Excel file for your city.**
  - a. If you are coding a new time period for a previously coded city:
    - i. Grab the city’s Excel file from the USD “City Files” folder. File name may vary slightly from the official city name.
  - b. If you are coding a new city never coded before:
    - i. Copy the “Template.xlsx” file.
    - ii. Rename the file to the name of your city.
    - iii. Set the CITY ID attribute in the “Overview” sheet to the city’s ID as you listed it in the “Cities” sheet of the “City List.xlsx” file.<sup>2</sup>

## What you need

1. Your City’s Excel file.
2. The “USD Coder” software (or just an ordinary internet browser if you prefer that).

---

<sup>1</sup> For coastal cities, use a Geographic Information Software (GIS) or get help from someone who knows GIS, to make sure the coordinates fall inside the city’s historical country border in the Cshapes dataset (<http://nils.weidmann.ws/projects/cshapes.html>).

<sup>2</sup> Make sure the ID in the “City List” file and the “[CityName]” file refers to the same city since the ID will be used to lookup the city and country information when building the final dataset.

## Searching for articles

1. Open the “USD Coder” software.
2. The software will ask you to login to your Keesings account. If you do not already have an account, creating one is free and easy to do on the [www.keesings.com](http://www.keesings.com) website.

---

*Using an internet browser instead?*

*Go to the Keesings website (<http://keesings.com>) and login there.*

---

3. In the search tab, search for the name of the city.<sup>3</sup> The from- and to- dates shall be set according to the time period that has not already been coded for the city you are working on. If the city name has a special character or accent, use the anglicized version of that character (e.g. write “e” instead of “é”). If there are multiple spelling varieties or historical names of the same city, separate searches must be made for each variation (e.g. first code all results for “Tehran”, then all results for “Teheran”, etc).

If instead you are continuing on a previous session, click the “Load” button and choose your city’s “.json” file that you saved in your previous session (or drag-and-drop the file into the application window). This should bring you to the article you were at last time and load the highlight terms you specified for your city.

---

*Using an internet browser instead?*

*Search using the Keesings website. The ordering shall be set to ‘Oldest First’ so that you code older entries first.*

---

4. Set up the Text Highlighter. By default the USD Coder program highlights different categories of words in different colors to help you highlight relevant unrest events. If you are starting a new search you must add the search terms to highlight the city’s country name in the “country” category, and the city name in the “city” category. The ‘\*’ wildcard character can be placed in the beginning, middle, or end of a word to allow for unknown variations in the beginning, middle, or ending of a word. The city category also looks for words like ‘capital’ or other famous city places like ‘Presidential Palace’, ‘Parliament’, etc., and you may add additional ones.

---

<sup>3</sup> The reason we only search for the city name and don’t add additional keywords or filter, is in order to make the search as exhaustive as possible and not exclude any events.

## Going through the reports

Screen each returned report individually. Screening a report involves three basic steps: looking for events, copying the relevant text, and codifying the event data.

1. **First, use the text highlights to help you look for disorder events in the city** that appear to be relevant according to the categories in the codebook. Note that if a report is entirely about your city, only looking for city name highlights is probably not enough and you should look closer at all mentions of unrest (e.g. if report was titled 'Major Unrest in Cairo' it is possible that the name Cairo only appears once).

When there are no highlighted colors in the "actions" and "actors" category panes, this indicates that there is no mention of disorder events in the entire article, so you can quickly skip to the next article. This is particularly useful for cities that generate mostly "peaceful" news about international relations, summits, economics, etc.

---

### *Using an internet browser instead?*

*When using an internet browser all keyword highlights have to be done manually one at a time (usually by pressing Ctrl+f). In this case it should be sufficient with a search of the city name and more detailed skimming of the text.*

---

2. **Second, copy and paste the relevant pieces of text into the Excel document's "Reports" sheet.**
  - 2.1. If you come across multiple events in the same report, each report-event should be given a separate row.
  - 2.2. Each row is given a unique REPORTID, consisting of a number that identifies the report (1, 2, 3, ...), followed by a unique identifying letter (A, B, C, ...) for the event within that report. Thus, a report may contain REPORTIDs 12A, 12B, and 12C, and another report may contain 76A and 76B. For some very long articles that list many events, you may end up reaching the letter "Z". In this case just prefix the ID with a letter starting at "A", and increment this the next time you run through the alphabet (A, B, C, ..., Z, AA, AB, AC, ..., AZ, BA, BB, BC, ...).
  - 2.3. Copy and paste the relevant text into the FULLTEXT field of the event/row that it belongs. Additional pieces of text for the same event should be separated by three asterisks (\*\*\*) to mark the discontinuity. It can be useful to make the Excel formula entry field bigger to more easily keep track of all the text.
  - 2.4. If it is difficult to determine relevance, the report text shall still be registered as a separate event and can be discussed at a project meeting. Report text mentioning an already coded event from a previous report shall also be registered if it provides significant additional information.

2.5. Once you have gone through the entire article:

2.5.1. Copy and paste the Keesings report title and date into the TITLE field. This should be the same for all events from the same report.

2.5.2. For each event found in the article, copy and paste a smaller subselection of the FULLTEXT field that more succinctly summarizes the event into the SUMMARY field.

3. **Third, codify the events in the “Events” sheet of the Excel file.**

3.A. If the event has not been coded before:

3.A.1. On a new line, select “Keep” from the STATUS field dropdown menu.

3.A.2. Increment the EVENTID with 1.

3.A.3. Code the various event variables in accordance with the guidelines in the codebook.

3.A.4. In the REPORTID1 field, please provide the event’s REPORTID described above (e.g. 1A).

3.B If the event has already been coded:

3.B.1. Change or update the event variables for the previously coded event if there is new information in the report.

3.B.2. Provide the report’s REPORTID in either the second or third REPORT ID fields.

4. **Remember to save your session regularly** in case something goes wrong, saving it with the name of your city, so you can easily continue where you left off.

## Finishing up

1. **When you have finished** going through all the search reports:
  - 1.1. **Register that you coded the city for the given time period in the “CodingHistory” sheet of the “City List.xls” file.**
  - 1.2. **Cross off the city as “Coded” in the “CandidateList” sheet.**
  - 1.3. If you have been working on a local copy of the city Excel file, **remember to replace/update this file in the “City Files” folder in Dropbox.**
2. **Making revisions.** If at a later point it is decided that an event should not be included in the dataset, only change its “STATUS” field to “Deleted”. This way we can better keep track of any changes and revisions.