*Project Goal*

The goal of the project is to digitize registrations from the Sommier Foncier, a French register that details transactions of properties and inheritances of these properties for the purpose of taxation. This is for scientific research on the return and risk to real estate investment, executed by a team of researchers from the University of Cambridge, Maastricht University and University Paris II.

The Sommier Foncier is a register available for entire Paris that contains separate sheets for each house, listing all relevant transactions. There are 3 series; the first serie (1809-1860), second serie (1860-1880) and the final series (1881-1943). There are some differences in the details across registers, but you will start digitizing from the first registers. The registers are organized per block of houses (part of a street).

*Data Entry Explanation*

Below we explain two pages of the first register of the sommier, and how to enter data from them. The file “EntryFile” contains examples of transcribed data from Register 26, up to photo 5617. The file “examples” help with transcribing the historical hand-writing.

The registers look the same across properties. The photos are in chronological order, but typically the right page is photographed before the left page. You start your analysis by looking at the top of the left and the right page.

On the top left of the left page it writes the type of property, all properties you enter should be “Maisons” if it says something else (e.g. “Terrain”) you should simply skip all photos from this property. Below it is the order number. The order number should be entered in Column C (“OrderNo”), and you enter it for this first observation of a property. Some properties span multiple pages, and in those cases the information on the top part of the page (e.g. the “Designation des Immeubles”) is often left blank on next pages. You can look at the order number to make sure you are still looking at the same property, since it appears on every left page. It typically increases one by one across properties.

The top of the right page contains a mention of the street and neighborhood and the property its house number. There is an old (“Ancien”) and a new (“nouveau”) house number. For the first observation of each house or street, you mention the name of the neighbourhood (Quartier, column D), street (Street, column E) and the **NEW** house number (HouseNo, Column F). Sometimes the new house number has been crossed out and written left or right of the original number: please make sure you identify the right one.

Next, you start entering information about the transactions. Often, the right page with rents has been photographed before the left page, so I start by explaining that page. You can also start with left page, as long as you make sure that all observations on one property are grouped together.

*Left page*

A screenshot of a computer

Description automatically generated

**Order No.**

*Right page*

A screen shot of a computer

Description automatically generated

The core of the right page contains information about rent prices. Observations on these can be on the entire page; often there is a “Bail” (rent) at the bottom of the page even if there is nothing written on top!

To start entering information about the “Bail”, you first look at the column “Objet des Locations”. You only enter an observation about a rent if the “Objet des Locations” is the **ENTIRE HOUSE**. This is often indicated as « maison », « maison entière », « maison en totalité », « ladite maison », « lade maison ». You do **NOT** enter information if this does not contain an entire house, e.g. if it writes « appartement », « appt », « divers lieux », « boutique », « rez-de-chausee » etc. For example, in the example photo on the previous page, none of the rents have to be entered. Note that the marks ‘’ are often used to indicate that description is the same as that in the line above it, so if you only see ‘’ you can assume the objet de location is the same as the objet de location just written above the marks.

If an observation contains an entire house you fill in the following information. In column G, “Type” you enter “Bail”. so we know this row is a rent price. In columns H, I and J you fill in the day, month and year of the rent contract, written in the column “Dates des Baux”. In the columns K, L and M you fill in the registration date, found in the column “Dates de leur Enregistrement”. In Column N, you fill in the length of the contract in years. Often there are multiple lengths e.g. 3, 6 or 9 years: if that is the case fill in the cell as 3\_6\_9.

Finally you enter the annual price of the property in Column O (Price). In most cases, there will be only one price. In some case, the price has multiple components, e.g. listing the “loyer”, “charges” and “impots” separately. In that case, you should add up all these numbers in Column O, so that it corresponds to the final price paid by the tenant. If there any important remarks relevant to the transaction, please enter them in Column Q.

Next, you can enter information from the left page, which contain information about property transactions. You look at the column “Nature des Mutations” and in Column G, you enter the “Type” of transaction. Nearly all transactions will either be “Succession” (inheritance), “Adjudication” (auction) or a “Vente” (regular sale). Very often, the transaction type is abbreviated or unclearly written. The Excel file “examples.xls” contains examples to help you. You write the full name of the type, also if it is not a succession, adjudication or vente. There are many possibilities. Sometimes, the sale or succession is only for part of a property. If that is the case, you write the part in Column P (Part). For example, if the Succession is for ½, you write 0.5 in Column P.

One special type of observation is the “Insuffisance”. An insuffisance belongs to the observation right above it in the register: if the previous record (e.g. a “Succession”) had a price that was too low, the insuffisance would correct that. If there in an insuffisance, please make sure if follows directly in the row below the observation where it belongs to; that way we can correct the price.

Next, you enter the data of the transaction and the registration, which are written under “Dates des Mutations” and “Date de l’Enregistrement ou de la Declaration”. You do so in columns H-M, similar to the rents. Finally, you enter the price in Column O. If there any important remarks relevant to the transaction, please enter them in Column Q.

*Remarks*

* If you are unsure if a rent corresponds to an entire property, do not enter that rent at all
* If you are unsure about a type of transaction, do not enter it at all and ask me a question about that observation in Upwork
* If you are unsure about a price observation, add a question mark instead of entering the observation.
* A second screen is not required, but highly recommended in order to be productive. You put the photos on one screen, and the excel file on the other screen.

*Way of working*

You will be in touch with the team via Upwork. Photos will be send digitally, either through a shared Dropbox, WeTransfer or other ways of communicating. GitHub might be used as well.

*Compensation*

The proposed compensation is 12 dollarcents per row of entered data, but you can also make your own price offer per row. In case of multiple qualified proposals, I will take the best-priced one. Note that one row corresponds to one transaction or rent. I am myself able to enter approximately 150 observations per hour.

For the first batch of data, you are expected to enter approximately 2100 rows of data. If we are happy with the quality and you with the job, we can continue with next batches. We have approximately 3500 dollars available for this project (we might have fewer photos than needed for this task, but we hope to take more photos).

*Apply*

Want to apply for the project? Please send the proposal via Upwork together with an Excel file where you make a small practice assignment. The practice assignment is to digitize the remaining photos on GitHub that are not yet in the file (the final photo digitized is image 5618). These are images 5619 to 5630. Please browse through the first set of photos and the digitized data first, to get an idea how and what needs to be digitized.

We plan to hire two to three people for this project, and hope to start within a few days.