Internship report - Summary

Research Assistant in the UMR LEDa-DIAL – IRD on an ANR project:

Transformation Of the French economy through the Lens of International Trade, 1716-1821 (TOFLIT18)

Quentin Vidal – TSE
# Table of contents

1) **Introduction** .......................................................................................................................... 3

2) **Organisms and project missions** ............................................................................................ 3
   a. LEDa and DIAL ...................................................................................................................... 3
   b. The TOFLIT18 project ......................................................................................................... 4
   c. Contents, advancement and aims .......................................................................................... 5
      Task 1: Collecting the data ..................................................................................................... 5
      Task 2: From data to research ............................................................................................... 5
      Task 3: Interpretative program .............................................................................................. 6
      Task 4: Going public .............................................................................................................. 6
   d. IT tools .................................................................................................................................. 6
   e. The datascape ....................................................................................................................... 8

3) **Works** .................................................................................................................................. 8
   a. Identification of commodities ............................................................................................... 8
   b. Transcript of sources ............................................................................................................ 9
   c. Orthographic normalization and merchandises’ simplification .......................................... 10
   d. Classification of goods ......................................................................................................... 11
   e. Meeting minutes .................................................................................................................. 12
   f. Writing of an article ............................................................................................................. 13

4) **Research Paper** .................................................................................................................... 13
   a. Introduction .......................................................................................................................... 14
   b. Data and method .................................................................................................................. 15
   c. Results ................................................................................................................................ 17
   d. Conclusion ........................................................................................................................... 19

5) **Final impressions and personal achievements** ...................................................................... 20

6) **Appendices** .......................................................................................................................... 22
1) Introduction

After four years of university education in the Toulouse School of Economics focused on macroeconomics, history and development economics, I now have to seriously consider my professional plans. This internship should constitute my first approach of the world of the academic research, a field that intrigues me as much as it interests me. Those three-month observation and work among a research team have offered me the opportunity to discover a multi-disciplinary project, involving historians and economists all connected by means of some useful, powerful and productive online tools. The conscientious building of such a large and essential dataset and its subsequent analysis have been very challenging aims; both of them require a great curiosity for this historical period, which perfectly matches with my interests. Through a variety of tasks and works, I have been able to add my own contribution to this project, and am thrilled that I can count it as such an excellent professional experience.

2) Organisms and project missions

a. LEDa and DIAL

The Laboratoire d’Économie de Dauphine (LEDa) was founded in 2009, and includes in 2015 about seventy five researchers and professors, as well as 120 PhD students and 6 administrative members. From then on, its main purpose has been to federate the different research teams in economics from the Université Paris-Dauphine around multifarious theoretical and practical projects, with a view to benefit from the association of complementary researchers in microeconomics and macroeconomics and to pool diverse economic tools such as econometrics, financial mathematics or even game theory. This research center is composed of five main thematic teams, which are:

- Financial Economics (SDFi)
- Energetic and Climate Change Economics (CGEMP)
- Health and Social Care Economics (LEGOS)
- Public Economics, Contract Theory and Game Theory (JEP)
- and the one I am part of: Development, Institutions and Globalization (DIAL-UMR IRD).
DIAL is a Joint Research Unit (called UMR in French) that gathers lecturers from Paris-Dauphine University and researchers from the Development Research Institute (IRD). It produces internationally renowned scientific work, provides academic courses of excellence in several countries and conducts surveys in developing countries, thanks to its partnerships and its locations abroad, for instance in Senegal and in Vietnam. It ranks as one of the main research centers in France in the field of Development Economics.

b. The TOFLIT18 project

From the beginning of my internship, I have been working with Guillaume Daudin, who is a full professor in economics in Paris-Dauphine University, member of LEDa (in the Joint Research Unit DIAL), and researcher in the Sciences Po Economics Research Center. He is also the coordinator of the ANR project TOFLIT18, which stands for “Transformation of the French Economy through the Lens of International Trade, 1716-1821”. This project was granted funding by the ANR (National Research Agency) in 2013, during four years and is supported by Sciences Po. Paris, the INED (National Institute of Demographic Studies) and Paris Dauphine. As Guillaume Daudin writes it, “Exploiting historical data to advance our knowledge in economy and history is obviously an inter-disciplinary task. All three partners are centers of research excellence and bring to the project economic historians used to working with both economists and historians. In addition, Paris-Dauphine brings its expertise in economics, the INED its knowledge of the treatment of large manuscript datasources and Sciences-Po its knowledge of dealing with databases through the Medialab.”

A wide variety of protagonists take part in the project, from Dauphine, INED, Sciences Po or other research centers. Besides Guillaume Daudin, I can mention Loïc Charles (INED coordinator), who is in charge of collecting the data and their transcript and forms the backbone of the project with my internship supervisor; Béatrice Dedinger (Sciences Po. coordinator), who is a researcher at the History Center of Sciences Po Paris; and Pierre Gervais, who is a lecturer in cotemporaneous history in the Sorbonne University. Here in Dauphine I am most of the time working with two fellow interns on the project: Corentin Ponton, who just finished his first year of Master’s degree in International Economics in Development in Dauphine, and Elisa Tirindelli, an Italian student coming from Paris School of Economics and achieving a master dissertation.
c. Contents, advancement and aims

The TOFLIT18 project aims at improving our knowledge of the French economy during the period that laid the economic ground for the entry of France and Europe in the modern industrial era; it is based on the transcription of French international trade registers in order to constitute a database as much exhaustive as possible on French trade statistics during the 18th century, as we will see later. Initially, it was divided into four stages (see the task schedule, Figure 1), but was a bit rescheduled through the months:

Task 1: Collecting the data

Guillaume Daudin and Loïc Charles have collected as many archives as they could in the Chambers of Commerce from many French cities such as Paris, Marseille, Bordeaux, Rouen and other coastal towns which hosted great harbors during the 18th century. They copied or photographed the trade registers or gathered the digital ones when possible (e.g. La Rochelle’s registers). From then on, a part of the team helped by some dedicated staff have transcribed the data on hundreds of excel files with a uniform “.csv” format in order to build up a dataset from the aggregation of those files.

Task 2: From data to research

Here, the team was aiming at designing and producing a private online datascape with a view to ease up the work of the researchers. The motivation was to allow members of the project to put the database in relation with similar external databases for other European countries (such as Great Britain and the Austrian Netherlands), and with maritime and merchant databases like MARPROF¹, NAVIGOCORPUS² and the Sound Toll Register³ databases. In this regard, the team started to encounter number of difficulties concerning the (ancient) French merchandises’ names, weights and measures, as we will see later. Anyway, the datascape is now online⁴ and does exhibit clear and nice tools which allow almost any user to explore the dataset we have built up, and to easily display plots and information on various trade flows regarding the indicators we want.

¹ MARPROF project: http://marprof.univ-paris1.fr
² NAVIGOCORPUS project: http://navigocorpus.org/
Task 3: Interpretative program

Though some delays have accumulated during the process, the team is now about to fully enter this stage. Some members even started to deliver papers and provide high scientific quality output, using both the datascape and personal data mining softwares such as Stata or R studio. Besides the most visible and obvious economic journals, they target more specific revues such as the Journal of Economic History, the Economic History Review, the European Review of Economic History and others.

Task 4: Going public

Here, the aim is to create a public version of the datascape that will be a simplified version of the private one, and will not interfere with intellectual property rights from other projects. It should allow both researchers to use the data and tools developed by the team and a broader public to access a clear visualization of significant aspects of the French economy during the 18th century. It is expected to be delivered during 2017, and in the meantime will be promoted by the publication of a book presenting the data. Most of the scientific production of the project, however, is destined to publication in peer-reviewed publications.

d. IT tools

During my internship, I have to use a bunch of very useful and intuitive IT tools, some that I was taught during my formation in TSE, and others that I was introduced to by my professor.

To begin with, the team decided to use LibreOffice Calc rather than Microsoft Excel for the comma-separated files (the “.csv” format) that they created during the transcript. They made this choice for two simple reasons: i) in order to harmonize the work form different members on the team, as LibreOffice Calc is a freeware, available on every operating system (Mac OS and Windows here), ii) because they wished to avoid encoding problems as much as possible, so they agreed to use the UTF-8 code on every single file of the project (so the French accents, often present in merchandise’s names, do not become Å© for example), and it appears that Calc easily permits to choose the encoding. I was more familiar with Excel but had already used LibreOffice for some past projects so it did not represent any issue.
Next, with a view to easily connect the works of the different members and to see the progress of everyone in the team, it has been decided to use the GitHub software, which provides a way to synchronize a whole folder of files online (much like Google Drive for instance), displaying moreover the different changes that have been made on the files every time a modification occurred. I have to admit that it caused me some trouble in the beginning, mainly because my computer does not seem to like the software very much each time I was launching a synchronization, but after some adjustments (I even upgraded my RAM), GitHub appears to be a very helpful and powerful tool once correctly handled (see Figure 2). Under some precautions, it can even synchronize the work of different persons on a same file in the same time, displaying the very changes made by each one of them and merging the whole in an “updated branch”.

At some points, my internship supervisor also took control of my computer through TeamViewer, which is a software program I had already used. In the beginning of my internship, it allowed him to get rid of most of the problems I was encountering with GitHub, directly from his personal computer. Additionally, we also made use of Livedrive, which is an online drive rent by the team to store nearly a terabyte of data, from photographs of files (all trade records, 18th century dictionaries, etc.), to classification rules and management methods, and of course every single database. By the way, I also had to install Notepad+ on my computer to deal with the huge “.csv” files (which are nothing but comma-separated text files); it offers a bunch of possibilities that Windows Notepad does not.

Furthermore, in order to coordinate every task assigned to team members, we use the Trello website, which is a kind of hub gathering the different missions each one of us is achieving. It consists in various tables, either private (for a member) or public (for the whole team), which display tasks as “To do”, “Doing” or “Done”, providing details and allowing everyone to comment each other’s tasks (see Figure 3). Trello is an intuitive and well-designed project management tool, and turns out to be very helpful for the whole team.

Finally, as I am going to explain more in the following, after two weeks I was given the possibility to start my own research work, in parallel with my daily tasks. Of course, I seized this great opportunity to catch a more precise glimpse of the world of academic research, and therefore I had to use software programs to understand and exploit the database
we are building up. In this regard I chose to work with **R Studio**, because I am used to this one, given that we were taught multiple R-Programming courses during the first year of master’s degree in TSE. Many members of the team prefer to deal with Stata, but I do not know anything of it and was told that both were very equivalent tools. Besides, R appears to be an excellent programing language when it comes to data mining and statistical analysis of datasets; so far, it allowed me to accomplish a correct (and needful) cleaning up of the data and to start some basic analysis in order to tackle the topic I am interested in.

**e. The datascape**

As mentioned earlier, one of the final targets of the project is the publication of an online platform allowing everyone (both academics and a broader public) to catch, quickly and easily, a nice understanding of the database. So far, the namely datascape is online but is only private, and is exclusively used by researchers for their interpretative works. Thanks to Paul Girard, an IT professional that has previously created a similar datascape on the RICardo project\(^5\) ([see Figures 4 and 5](#)), we can intuitively explore the metadata, so as to realize where and when data are present or missing. Besides, we are currently able to draw roughly any basic plot we want, by the means of different indicators ([see Figure 6](#)). At the end of my internship, the platform should be largely improved and we would like it to be soon public.

**3) Works**

**a. Identification of commodities**

At the very beginning of my internship, the first task Guillaume Daudin assigned me – after having downloaded and signed in to the list of pieces of software and websites I mentioned earlier – was the identification of a number of commodities, for they were causing us some concerns. Indeed, after spending approximately two years on the transcript of trade registers and national resumes about the French imports and exports during the 18\(^{th}\) century

---

\(^{5}\) **RICardo** is a project started in 2004 by Béatrice Dedinger and Paul Girard that deals with the whole set of bilateral fluxes which composed the International Trade from 1800 to 1938. It is therefore an analogue project of TOFLIT18, except that it is not centered on 18\(^{th}\) century France. The datascape is available online here, without any password: [http://ricardo.medialab.sciences-po.fr/#/](http://ricardo.medialab.sciences-po.fr/#/).
on LibreOffice Calc, the team built up a list of merchandises’ names (roughly twelve thousands once corrected and simplified, amongst more than forty thousand commodities in the whole date base) that they did not know, because they could not see what they mean. Therefore, following a precise method, we (me and Corentin Ponton, the second intern that joined the project two weeks after me) have dealt with every problematic word. For a large part of them, it appears that it was simply an orthographic problem: the transcripters have obviously committed typos here and there when entering pages and pages of manuscript trade registers into “.csv” files (see Figures 7 and 8). Sometimes they were easy to spot, but most of the time we had to check the photograph of the appropriate file on Livedrive (an online drive rent by the team), to see whether it was indeed a typo, or if on contrary the transcript was totally well-done. In such cases, we had to pass on the second stage of the identification: perhaps the word that was causing us trouble was correctly written, and therefore it simply means that we did not know what it referred to. However, because the final aim was to provide a clear and understandable database for everyone, there must not remain any undefined words. To correct this, we were given a handful of dictionaries and encyclopedias from the 18th century, because the French tongue has evolved a lot since, and particular spelling might have changed or basically disappeared. Amongst them, we can mention The Savary, The Littré, The Diderot and D’Alembert’s Encyclopedia, or even more specific dictionaries, such as ones focusing on trade goods, arts and industry; they are all available online, either on Google Books, or on the personal LiveDrive of the team. Then, if we were to find a definition that matched our word, we simply reported it in our list, and marked the word as “defined”. If we were not able to define it however, we marked it undefined, and let the unknown merchandise to further analysis: perhaps it was written in a foreign language (often old Flemish, or local dialect), perhaps the 18th century guy who created the trade register committed a typo.

I must admit it was not the most exciting task I have been doing so far but it did consist in a good introduction to the project, and made me understand how meticulous the building up of such a dataset was.

b. Transcript of sources

Once we finished the identification of commodities, we were asked to transcript some new sources into LibreOffice Calc files. As explained before, it consists in the conversion of
trade registers from the paper format to a numeric format, here the comma-separated format. It is a huge part of the project; certainly not the most fascinating one but still the basis of every work we aim to do. Although the large part of trade reports and business registers have already been copied and converted into “.csv” files during the first years of the project, there were still some photographs in “.pdf” format that were not exploited and could enlarge the size of the global dataset. The different sources created from those documents are then grouped by origin (local, national, colonial space, etc.) and then aggregated in the final dataset. As a matter of fact, I was given a number of these registers to transcript, captured by team members in the dawn of the project in various French Chambers of Trade across the country (or downloaded online on prefecture’s website when this was possible\(^6\)). Despite the trade registers largely vary by size and content, mines consist in documents, each of about fifteen pages per year, which recall all the French international trade that has been made during this year, classified by categories (mineral and vegetal matters, chemical products, animals, etc.) and displaying, for every flow, the price, the quantity traded, the unity of measure, the unit price, the value it represents and the custom duties paid (both in Francs, see Figure 9 and 10). Moreover, indications are sometimes given concerning the provenance or the destination, and the particular rules regarding French colonies and fees exemptions. Luckily enough, the “.pdf” files I had to transcript were photographs of post-1815 typed registers, instead of the classic handwritten documents which often prove very difficult to read. So far, I have entered about half the files I was asked to do and will partially continue this job for some weeks, along with the research work I started earlier in May and that I am going to talk about later in this report.

**c. Orthographic normalization and merchandises’ simplification**

While team members were transcripting a large quantity of trade registers from different sources in order to build up the dataset, they face a critical issue: a same product, whether it is reported in a local La Rochelle’s trade register or in a Paris’ national resume of French trade, can be written differently. Sometimes the appellation of the merchandise differs from a region to another (nowadays for example we know the famous case of “chocolatine” in South France versus “pain au chocolat” in North, but the matter was largely more frequent in

\(^6\) See La Rochelle’s registers on the prefecture website, labelled as “41 ETP 270” and “41 ETP 271” documents.
those days); sometimes the spelling differs, as for the French word for “scythe”, which can be written “faux”, “faulx”, “fauls”, “faulse” and so on… Indeed, the 18th century French orthography was much less standardized and established that it is today. To tackle this inconvenient issue, team members decided to accomplish an orthographic normalization, subsequent to a simplification of the goods’ denomination, by following a set of pre-established rules. Details of this complex process are available on the TOFLIT datascape, but basically it firstly consists in normalizing the orthography of each merchandise’ occurrence (in my previous example, all the words referring to scythe were gathered under the set spelling of ‘faux”, because it is the most frequent spelling or because it is the current spelling); and secondly in grouping them by meaning, so as the French word “chocolatine” and “pain au chocolat” can be understood as the same commodity. The first step reduced the number of items from 45 473 to 17 002 (!), whereas the second step gathers those 17 002 well-spelled products into 11 387 groups of product referring to the same things (see Figure 11). Although some normalization and simplification has been done automatically thanks to Stata scripts (by simply removing plural forms for instance, or every upper-case letter), the great part of it must be manually treated, by the means of ancient dictionaries (approximately the same ones that I mentioned for the identification of commodities). Furthermore, it is not yet a completed task, as we are still facing issues identifying merchandises. As explained by a previous intern on the project, we happen to find inconsistencies when cross-checking local sources (coming from ports) which, once aggregated, should exhibits the same amounts and values of trade than national sources (coming from the National Office of Trade). Not taking into account the frequent computing errors made by 18th century copyists, we understood that each one of them was not accounting goods the same way than others. He mentioned the case of French Normandie in the 18th century which was a great producer of canvas; in this time, every town made their own specific canvas, and while a local copyist who registered every trade flow was able to distinguish a canvas of Elbeuf from a canvas of Rouen (and so to enter them differently in the record), a national copyist would not have made the difference, and would have simply entered “canvas of Normandie”… Those many little inaccuracies make a great difference once aggregated in a 400 thousand observations dataset, hence we need to correct them as much as we can in order to relevantly exploit our data.

d. Classification of goods
After the orthographic normalization and simplification, what one can call a third step in the treatment of commodities is the classification. With a view to make the final dataset public at the end of the project, Guillaume Daudin would like every user – and even someone who does not know anything in International Economics – to be able to understand and handle the dataset on the datascape. In this regard, the team decided to classify every single commodity according to the Standard International Trade Classification (SITC Rev.3), which is a global and well-known classification proposed by the United Nations. It is available online⁷, albeit the version we are using has been a bit adapted to serve the purpose of a project dealing with goods of the 18th century. It groups products under 9 main categories, themselves composed of many sub-categories (category 1 gathers “beverages and tobacco”, category 2 “crude materials, inedible, except fuels”, and so on). Besides, another key advantage of this classification is that it allows creating bridges between different projects on international trade, whatever their epoch and their region. I was not directly involved in this task (Pierre Gervais is), but sometime I had to use it when I was identifying commodities, or when I am doing my research work to isolate some precise kind of goods as I am going to explain later.

e. Meeting minutes

Every month, a TOFLIT meeting is held. So far I have attended two of them, in the MediaLab of Sciences Po Paris. All members of the project and interns are conveyed to discuss the global progress of the project, to emit some remarks or basically to share what went ok and what did not since the last reunion. Everyone takes the floor in turns, and generally the meeting is concluded with a numeric presentation of the advancement made on the datascape, and with a recapitulation of new tasks to assign in near future regarding the scheduled agenda. What is great in this kind of reunion is the presence of researchers from different horizons: some are economists, others historians or even computer scientists, and every one of them brings its own specific contribution, learning from each other’s aptitudes. I was asked to write meeting minutes that were uploaded online after these gatherings, in order to debrief what was said and to inform any researcher who could not come on the main themes covered. Moreover, every year is held a bigger meeting, where different guests are conveyed to discuss and relate the TOFLIT18 project with other projects, to cross-check difficulties encountered.

⁷ SITC rules as established by UN: http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=14
with 18th century data or to have an insight of what kind of dataset could be published once achieved. The 2016 reunion will take place in Chicago, in October, with many international guests.

f. Writing of an article

Besides doing transcript and research work, Corentin Ponton and I were also offered the possibility to write a 1000 words article introducing the TOFLIT18 project, emphasizing on the datascape. Thanks to other articles, and helped by Guillaume Daudin and Paul Girard, we quickly wrote a first draft, which was immediately proposed to Sandrine Mesplé-Somps, headmistress of the Joint Research Unit DIAL-IRD-Paris Dauphine. As a matter of fact, firstly it was only designed to be a newsletter for DIALogue, a bi-yearly review published by DIAL. After a lot of corrections and revisions, it was finally published online on the DIAL website.

However, we also decided to propose this article to TheConversation website, an online independent journal that publishes academic views and research works about multifarious fields, especially economics. We are currently waiting for their answer.

4) Research Paper

In May, Guillaume Daudin offered me the possibility to start my own research work on the project. Obviously, I seized this great opportunity with a view to better understand the different aspects of the job of researcher. Thus, since this time I have spent approximately half working days on this task, while the other half was dedicated to identifying commodities in the first place and to transcript sources afterwards. My internship supervisor did not give me any bonds regarding the topic I would like to cover, providing of course that I made use of the TOFLIT18 dataset we were building up. After some days of hesitation, I decided to work on the First French Colonial Empire and more precisely on its dislocation in the second half of the 18th century. Especially, I resolved myself to approach the case of the Seven Years War.

8 Link: http://www.dial.ird.fr/media/ird-sites-d-unites-de-recherche/dial/documents/publications/dialogue/dialogue-43-focus
with England and its effect on the French economy, through the study case of the North American trade. Indeed, the Treaty of Paris that ended this war in 1763 consisted in a huge humiliation for France (and Spain), seeing that it imposed the former to yield a great part of its colonial empire to Great Britain, allowing the world to enter a period of British dominance. Following this forced peace, France had to cede – amongst other concessions – Canada, Dominica, Grenada, Saint Vincent and the Grenadines, and Tobago to Britain, as well as the eastern half of French Louisiana. In my work, I decided to focus on the loss of the North American territories subsequent to this war; I have tried to understand to what extend the French colonial trade was affected by this defeat, and to see if France eventually found a way to substitute North American products by others or if it simply diverted its imports.

a. Introduction

From 1689 to 1815, England and France competed in the so-called “Second Hundred Years War”, through a harsh commercial rivalry. To Linda Colley, it consisted in “a world fight between France and Great Britain for the imperial and commercial supremacy”. However, during what might be called an “extended 18th century”, years of military peace were largely dominant, interrupted here and there by fights, such as the Seven Years War which started in 1756 and ended in 1763. This conflict is considered as the very first World War, given that it involved most of the great powers of the time (Great Britain, Prussia, Portugal and allies against France, the Holy Roman Empire, Austria, Spain, Russia and allies) on various places of the world. Nevertheless, it is also considered as a commercial war; Richard Pares even said that the Seven Years War was nothing but a “competition over a group of sugar colonies and another one”.

In 1693 already, Josuah Child affirmed in a very mercantilist fashion that “all trade is a kind of war”, stating that a country could only get rich to the expanse of others. In this regard, while Colbert tried and launched an industrialization policy in France, England chose to develop its Navy, to finally make it the most powerful armed force overseas. By controlling the seas, Great Britain controlled the colonial trade, and vice versa: by getting richer, it became stronger.
In 1738, worried by the French renewed strength on economic and military plans – mostly permitted by the soaring of the production in the Antilles –, the Common Sense wrote that “trade that rises can be destroyed by war; trade that lowers gains a chance to live back thanks to it”. According to François Crouzet in his book “La guerre économique franco-anglaise au XVIIIe siècle”, English pamphleteers actually feared peace time with Spain and France, arguing that nothing but a war could improve their economic situation, by lessening others’ armies and above all trade.

b. Data and method

Despite the careful treatment the database benefits after the crude transcripts and aggregation of French trade records (such as a detailed orthographic normalization preceding a massive simplification of goods and countries’ names with a view to classify any merchandise according to the SITC rules), the database still needs mending. Added to the basic cleaning up I had to apply in order to correctly perform any statistical analysis with R Studio, it required special treatments.

First of all, small modifications were needed: I for instance transformed revolutionary years that appear after the French Revolution (Year 2, Year 3, and so on after 1992) into date on a standard Gregorian format, as well as I removed particular observations that did not offer any supplementary information concerning my purpose (custom fees and tariff for instance).

As a second step, I decided to constitute a subset of the whole dataset, because I wanted to work only with the trade of France with its North American colonies, namely the Nouvelle France. Here I encountered my first main issue: indeed, the words “Canada” or “Louisiana” for example are very rarely mentioned in the trade records of the 18th century; copyists were limiting themselves to destinations and provenances such as “French colonies” or “American islands”, but never more precise. Therefore, I have proceeded differently to work around this problem. I firstly created a subset of products exclusively imported from the French colonies. Out of the initial 400 thousands flows (counted as rows of observations) composing the global database, about 40 thousands were exchanged with French Colonies, and only 6 thousands were imported. Then, when it was detailed enough, I removed the lines of trade that were made with other entities than North American colonies. As a matter of fact,
copyist sometimes added precisions about the origin of commodities, especially when they came from Antilles (Guadeloupe, Martinique and so on). Once done, 5500 observations remained as “possibly imported from North America”. Later, I collected every merchandise name in a table in order to analyze them: I got about 700 unique items (copper, leather, medicinal drugs, iron, flour, wine, etc.). I carefully examined each one of them and marked them differently thanks to a dummy called “provenance” according to four levels: “exclusively North America” (e.g.: beaver leather), “potentially exclusively North America” (cod fish), “non-exclusively North America” (iron) and “intractable”. In the end, I found about 70 unique commodities that could only come from North America. Hence I merged my results with the full dataset and was now able to track the French trade with North America.

Finally, 707 lines of trade concern those “exclusively North American” commodities in my sub-sample of colonial imports. They go from 1718 to 1779 and are mainly trade with Quebec and Louisiana. As can be expected, they largely concern a few groups of the SITC classification (Figure 2), because they are mostly “plantation foodstuff” (0b) or “leather products except saddlery” (6a).

**Figure 1:**

![SITC Classification](image)

Nevertheless, before any econometrical analysis could be performed, I had to apply the “repeated values” process, as Elisa Tirindelli and Pierre Hollegien did in their work. Indeed, though the whole dataset is more than 400 thousand rows long, we still miss information regarding particular years or places. For instance, although we collected every yearly trade records emerging from La Rochelle since 1718 to 1779, we could not find any registers reporting the trade in Marseille after 1756. Thus, we have to estimate trade values, assuming that no major structural changes happened during this period. Basically, it consists
in predicting the trade for i) directions (Marseille, Rouen, Bordeaux, etc.) and ii) years that we do not know about, according to data we have for other directions and years. Furthermore, one might want to point out that this treatment is run with aggregated values of trade per year and per direction, and not with every single line of observations. Once those total sums of trade computed, I created two set of dummies, date (t) and direction (i), and took the logarithm of total sums of imports in order to run the following linear regression:

$$\log(\text{imports}_{ti}) = \beta_0 + B_1 \text{year}_t + \beta_2 \text{direction}_i$$

This model aims at predicting the aggregate imports per year and direction regarding the yearly changes in trade composition by source, with the assumption that the composition is constant trough time. In the end, I obtained a set of regression coefficients (one for each year and for each direction), which allowed me to predict the trade in every missing case. For example, though we do not dispose of any register reporting the trade in Nantes in 1750, we can assume that it is:

$$\log(\text{imports}_{1750,\text{Nantes}}) = \beta_0 + B_{1750} \text{year}_{1750} + \beta_{\text{Nantes}} \text{direction}_{\text{Nantes}}$$

$$\iff$$

$$\log(\text{imports}_{1750,\text{Nantes}}) = 13.2585 + 2.1217 \times 1 - 3.3216 \times 1 = 12.0569$$

$$\iff$$

$$\text{imports}_{1750,\text{Nantes}} = 172284.1$$

Overall, predicted values fit well with what can be expected regarding actual total sum of trade for known directions and dates.

c. Results

From then on, I was able to conduct a small econometric analysis. To begin with, I basically drew a simple graph that exhibits the evolution of French imports of the North American commodities through years (Figure 1). It is important to note that it concerns only the trade made with French colonies (according to the list I had to make previously), and only commodities that exclusively come from North America.
As can be expected, the trade fluctuates a lot before the Seven Years Wars, reaching a peak in 1741 with an equivalent of 27 208 911 Tours pounds (i.e. Livres Tournois). Although trade also knew years of very low activity, in 1726 and 1929 for instance, it was roughly constant and important through this period. During and after the Seven Years War however, and the subsequent signature in 1763 of the Treaty of Paris that forced France to cede most of its North American colonies to England, one can clearly distinguish a breakpoint. Imports have completely plummeted and stagnated around 0 from then on. To confirm the visual guess, I decided to perform a Chow Test around the year 1763. Through R, I divided my dataset into two subsamples around this critical year and run the procedure. It is completely conclusive, seeing that the p-value of the test is extremely low (0.017).

Despite what we could think, it does not exactly mean that France stopped to import North American products, but rather that it stopped to import from its North American colonies. Indeed, once under Britain control, colonies were subject to trade ban with France and other European powers as a measure of retaliation. Recall that Great Britain aimed at ruling overseas, by military and commercial means. Therefore, although the French imports of North American products directly coming from North America dropped at the end of the war, one might want to conduct an extensive analysis, tracking those very products with a view to
find out how France managed to divert its imports afterwards, whether by substituting North American products with equivalent from other places in the world, or by keeping buying them through England. One way to process could be to use the list of commodities that were exclusively imported from North America (back when France disposed of colonies there), to track those specific products in the whole dataset (checking again that they truly cannot be produced anywhere else) and to see by which channels they were brought (if they were) to France after the Treaty of Paris. For instance, one could approach the case of the castor leather sold to France by England during the 18th century; if this commodity trade was to increase by a substantial amount after the Seven Year Wars, then it would mean that France did not stop consuming North American products, but that it now has to buy them directly from England. On the other hand, it could be interesting to see if France substituted North American products by available equivalents, especially for leathers, furs and skins.

d. Conclusion

According to James C. Riley in his book “The Seven Years War and the Old Regime in France – The Economic and Financial Toll”, although the Treaty of Paris in 1763 consisted in a great humiliation for France, it remained the first European power for a while. Riley shows by counterfactual that, thanks to good anticipations and a wise management, the quantities traded during the war have not been much less important than what they could have been, had the war never happened. Because Britain cut trade route overseas and eventually prevented France from bargaining with English colonies (both the old and the newly acquired ones), France had to swift its international trade. It performed a needful deviation toward continental Europe under Napoléon and substituted specific products from Nouvelle France by new ones (respectively by the means of blockade and import-substitution industries). Although the Eden Treaty of 1786 was designed to improve commercial terms by pushing for trade liberalization and cutting taxes, with a final view to put an end to the hundred years’ commercial war between the two great powers, it met a quick end. Indeed, the French Revolution occurred, accompanied by an increase in French tariffs between 1791 and 1793 (Azam); eventually, protectionism ruled Europe.
5) Final impressions and personal achievements

After three months of internship, this professional experience proved itself very enlightening in different aspects. The main one was my principal motivation while I was looking for an internship: I could finally catch a glimpse of the world of the academic research. For what I can say of it, I find it very specific, very distinctive from other fields of work a student in economics could consider. In my opinion, research requires i) a great personal autonomy and ii) an insatiable curiosity, for your personal preferred field of course but also for every field of studies that could serve your work. While I think that I perfectly fulfill the later requirement (I have always been curious and enthusiastic about any new topics of study), this internship has been giving me some huge skills in the former. I was quite free and independent here; most of the time, Corentin Ponton and I were working on our own in an office provided by the laboratory, close to Guillaume Daudin’s one. We did not have precise delays to respect as one could expect in other companies or institutions, however we were always working on something, but rather because we got to be interested in it than because we were asked to.

Furthermore, this internship offered me the possibility to work on a multi-disciplinary project, involving economists of course, but also historians, data scientists, web designer, research director and so on. They all share their skills and knowledge with others, providing a lot of different angles to approach the topic of the French economy in the 18th century. Once or twice a week, we were also given the opportunity to attend scientific seminars: they are dispensed by various researchers from Dauphine and other universities like Paris School of Economics or Stanford, but also from other international organisms like the Research and Development Institute (IRD) or the Banque de France.

Nevertheless, while the topic tackled by the team was very interesting from a personal point of view, I am not sure it convinced me to work in the academic research. Patience, considerable personal investment and ability to succeed in a quite competitive environment are all required. Maybe I am not passionate enough about a particular subject to make it my daily bread and butter, maybe the ten-year schooling discourages me. On the other hand, working on a subject that completely matches my interests has never been a concern to me, thus it could simply be a matter of time before I find the perfect field and devote myself to it.
Nonetheless I was very glad and thankful to be part of this team, to modestly contribute in the advancement of an ANR project and to draw from it valuable learnings and humble pieces of work. I am really looking forward to see what will comes next in this research project and will obviously keep an eye on the progress of the study, through the datascape or thanks to newsletters.
6) Appendices

Figure 1: Task Schedule
Figure 2: GitHub interface

Figure 3: Trello interface
Figure 4: RICardo - Graph reporting entitles table, that is, providing detailed information on the amount of available data for each countries and years.

Figure 5: RICardo datascape - Plot reporting the total trade (in sterling pound) between United Kingdom and France, from 1831 to 1938. One can clearly spot the WWI breakpoint which caused a lack of data.
Figure 6: TOFLIT18 datascape - View by network of import flows between directions (●) and countries (●); 1716-1821.
Figure 7: “Tableau Général”, National register to transcript, 1860.

Figure 8: The corresponding comma-separated file on LibreOffice Calc.
Figure 9: “Table of Quantities”, National Register to transcript, 1822.

![Table of Quantities](image1.png)

Figure 10: The corresponding comma-separated file on LibreOffice Calc.

![Comma-separated File](image2.png)
Figure 11: Views of the orthographic normalization and merchandises’ classification (available on TOFLIT18’s datascape).

Figure 12: Plot on R – Evolution of French trade with any destinations (here French Colonies)