

# FAQ Aziende App

## What is it?

Aziende App is a tool designed to facilitate the search of companies in Emilia Romagna. With Aziende App, you can easily filter enterprises according to different criteria you define. It has two sections. The filter section filters the database according to the criteria you define. The cluster section returns similar companies to the company you ask for.

## What database is being used?

The database used by the app was downloaded from the AIDA website in February 2021. It contains 16428 companies, which are all the companies from ER which at the time of download were active, not bankrupt and had 10 or more employees. It contains the following variables:

- Identification: Ragione sociale,
- Location: Provincia, Comune, Codice postale, Indirizzo, Website, Numero di telefono
- Size: Ricavi delle vendite (nmil EUR), Utile Netto (nmigl EUR), Dipendenti
- Description: Settore, Forma giuridica

## How to use it?

Follow this link: <https://rubenmartinezcuella.shinyapps.io/AziendeER/>

It is optimized for the desktop, not for a phone. Keep in mind that the server of this website has a limit of 25 hours per month for its free version. This should be enough for normal usage, just remember to close the window once your search is done.

You are presented with “sezione filter” and “sezione cluster”. Click where you want to go.

**Sezione filter:** you have the filter criteria on the left. The default values are the less restrictive filters, so if you click “Go!” without any selection, all companies will be displayed. The filters are self-explanatory, there is only one remark to be made. “Settore industriale” allows you to choose from 20 categories that cover the entire market spectrum. “Area specifica del settore” further displays subcategories from the selected sectors of the previous 20. Note that this is a more restrictive filter: if you select all in “Settore industriale” but then select just one subcategory of those displayed in “Area specifica del settore”, the companies displayed will only belong to the specified subsector. If you don’t wish to differentiate in subsectors, either select all in “Area specifica del settore” or leave it blank.

<div>Settore industriale</div> <div>20 items selected</div>	<div>Settore industriale</div> <div>20 items selected</div>	<div>Settore industriale</div> <div>20 items selected</div>
<div>(Non è obbligatorio modificarlo)</div> <div>Area specifica del settore</div> <div>ESTRAZIONE DI MINERALI METALLIFERI</div>	<div>(Non è obbligatorio modificarlo)</div> <div>Area specifica del settore</div> <div>88 items selected</div>	<div>(Non è obbligatorio modificarlo)</div> <div>Area specifica del settore</div> <div>Nothing selected</div>

As a practical example, the image on the left will only display the companies involved in “estrazione di minerali metalliferi” despite having selected all 20 sectors, while the one on the centre and on the right are equivalent, and will display companies involved in all the sectors.

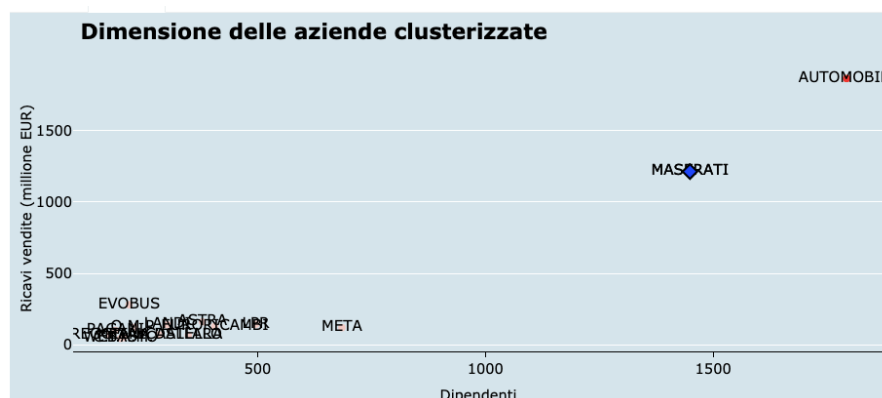
Once you click “Go!”, below you have 3 buttons. With the first, you will observe on the map the physical location of the headquarters of the companies filtered. The second, “Grafico”, shows a scatterplot of the turnover and employees of the filtered companies. This can be helpful to understand your selection in general terms. Finally, the “Dataset” button contains the variables defined above. Note that you there is a search box, which you can use to find a specific company by its name.

Click on “Scaricare i dati” to download your selection and on “Help” to obtain some more information.

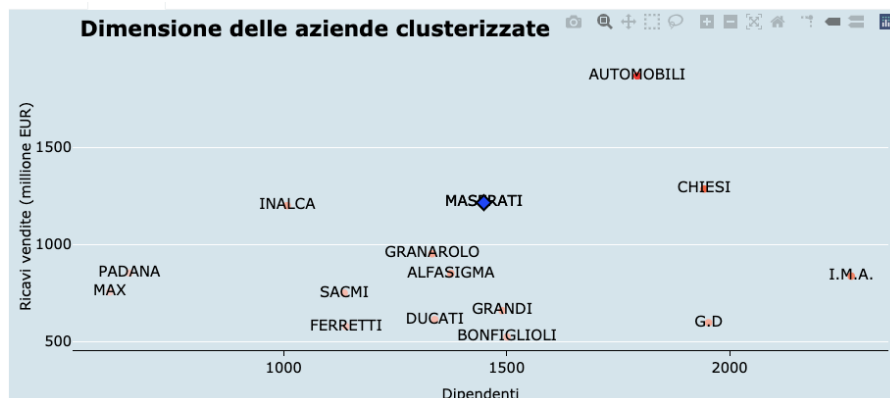
**Sezione cluter:** you supply the codice fiscale of the company you want to find similar companies to. It then finds the most similar companies in terms of revenue and number of employees (with equal weight) given the constraints imposed on province and sector. If you don’t know the codice fiscale, go to “Sezione filter”, click “Go!” without any selection, click on “Database” and type in the search box the name of your company. Then copy paste it.

The selection of “Volete che la corrispondenza sul settore sia ampia o ristretta?” is critical. There are three levels. Using Ferrari as an example, “ampia” will consider all consider all companies in “ATTIVITA MANUFATTURIERE” (it considers its “Settore Industriale”). This, for example, also includes food industries. The next level, “intermedia”, would consider all companies in “FABBRICAZIONE DI AUTOVEICOLI, RIMORCHI E SEMIRIMORCHI”, this is, its “Area specifica del settore”. Finally, “ristretta” would limit the clustering to enterprises in “FABBRICAZIONE DI AUTOVEICOLI”. This can get very specific in certain cases.

Finally, impose restrictions on province, select the number of most similar companies you want to retrieve (sometimes there is a -1/+1 error, I could not fix that bug) and click “Go!”. The tab “Grafico” is especially important here. Your company will be displayed in blue, and the others in red. If your company y more or less in the centre, it is a good symptoms, it means there are indeed similar companies. Big companies like Ferrari are natural outliers, so don’t be surprised to not find it perfectly in the centre.



*This selection for Maserati is not great. The “closest” companies aren’t that close. If this happens, try to restrict less your selection i.e. set corrispondenza to ampia.*



We observe how allowing for other companies not only in automobile industry places Maserati in the middle, meaning those companies are more similar

The other functionalities are equivalent to the filter section.

## Advanced/technical information

**Running it in your local server:** you can run the application for unlimited time in your local server. Just make sure you have R installed and the dependencies installed from the beginning of the file. Find the code here: [https://github.com/rubenmartinez9/Search-for-companies-in-Emilia-Romagna/blob/main/Aziende\\_app.R](https://github.com/rubenmartinez9/Search-for-companies-in-Emilia-Romagna/blob/main/Aziende_app.R)

**Update database:** this part is a bit more tricky, but if you know some R, you can manage if you follow these recommendations.

1. Go to AIDA and select the criteria of the companies you are interested in (for example, you may want to include companies with under 10 employees).
2. Download the same variables as in the previous database. This is:

**La tua selezione:**

- ☒ Variabili
- ☒ Ricavi delle vendite (Ultimo anno) | migl EUR [Modifica](#)
- ☒ Dipendenti (Ultimo anno)
- ☒ Codice fiscale
- ☒ Sede operativa - Comune
- ☒ Sede operativa - Provincia
- ☒ Sede operativa - Regione - Regione
- ☒ Indirizzo sede legale
- ☒ Sede operativa - Codice postale
- ☒ Sede operativa - Numero di telefono
- ☒ Website
- ☒ Forma giuridica
- ☒ Descrizione attività (IT)
- ☒ Attività principale
- ☒ Principali prodotti e servizi
- ☒ Start-up innovativa
- ☒ Utile netto (Ultimo anno) | migl EUR [Modifica](#)
- ☒ ATECO 2007 codice
- ☒ ATECO 2007 descrizione
- ☒ Indirizzo sede legale - Longitudine
- ☒ Indirizzo sede legale - Latitudine

Now you have the database. Let's integrate it with the app.

3. Create the 3 levels of sector from the ATECO 2007 codice. The 3<sup>rd</sup> one we already have it, it is the variable ATECO 2007 descrizione. Unfortunately I have lost the code to create the first 2 levels, but it shouldn't be too hard. Just extract the 2 first digits of the ATECO code and merge it with this excel, for which we have level 1 and level 2. [https://github.com/rubenmartinez9/Search-for-comanies-in-Emilia-Romagna/blob/main/sector\\_codes.xlsx](https://github.com/rubenmartinez9/Search-for-comanies-in-Emilia-Romagna/blob/main/sector_codes.xlsx)  
Now, we have all the above variables plus the 3 levels.
4. We add the variable "region".  
This may sound confusing because we already downloaded the variable "Regione". However, we need to make a few changes because the package that maps the companies doesn't know Rimini exists, and it calls other provinces with slightly different names. The code is here: <https://github.com/rubenmartinez9/Search-for-comanies-in-Emilia-Romagna/blob/main/adjust%20region%20from%20AIDA.R>. You will probably have to adjust the code a bit, or rename your variables.
5. Now we have the entire updated dataset. If you pass this test, the new dataset will work. Check if the variable names are the same as the previous dataset, which you can find here: [https://github.com/rubenmartinez9/Search-for-comanies-in-Emilia-Romagna/blob/main/ER\\_V3.csv](https://github.com/rubenmartinez9/Search-for-comanies-in-Emilia-Romagna/blob/main/ER_V3.csv). If yes, BINGO!, it will work. Change line 15 of the code of [Aziende App.R](#) to add the new address to load the updated dataset.
6. At this point you can run it in your local server. Unfortunately, you can't access the current online server shiny, since it is deployed from my computer. You have two options if you want to make it public like mine: a) Create a new account on <https://www.shinyapps.io/admin/> (recommended) or b) Send me an email with the updated code and I will deploy it from my computer ([rubenaustralia@gmail.com](mailto:rubenaustralia@gmail.com)).

**Want to upgrade the app?** if you are reading here, I reckon you must have experience on R. There are a few things that can be upgraded by modifying the code.

- Create a section with time series financial data. You can download from AIDA the previous 8 or 10 years, so it could be a good figure to compare. This way, you could filter/cluster by growth as well.
- Add a section to explore certain selections. So far, the filter just provides the companies. It could be interesting to write a short program that computed some summary statistics of the selections. This way, you could probably explore different selections to try to find e.g. a specific sector that has had a huge growth in the last 2 years.