



Europass Curriculum Vitae



Personal information

| | |
|------------------------|-----------------------------------------------|
| First name / Surname | Andrea Melloncelli |
| Address | Via Contardo Ferrini, 81, Samarate (VA) 21017 |
| IT VAT Number (P. Iva) | 03494930120 |
| Telephone | +39 348 74 32 874 |
| E-mail | andrea.melloncelli@gmail.com |
| Skype | andreamelloncelli |
| Nationality | Italian |
| Date of birth | 12/05/1983 |
| Gender | Male |

Summary of experiences

I work as a freelancer Data Scientist Consultant and Trainer.

I teach Statistics and how to use statistical software to solve business problems, my expertise range from calculus sheets (Excel and Google Sheets) to programming software as Python and R, depending on the type of problem to solve. I recently taught in Milano Bicocca and Milano Cattolica and I have experience of training on the job in several companies.

In 2019 I have co-founded VanLog, a team of data scientists with a strong focus: providing full stack data science solutions particularly suited for small and medium business. All the tools that are needed to understand and communicate data are part of our business, from statistical modeling to business intelligence techniques, ad-hoc programming, report automation, creation of interactive dashboards, integration via API.

As a qualified physicist with a passion for Data Analysis and new Information Technology, I have a deep and solid experience in R, Shiny programming and development along with extensive skills in unix system management, IT automation tools, cloud technologies, big-data platforms such as Hadoop and Spark.

Previously worked for Quantide Srl (since 2015) as an IT Management Consultant and "R" Trainer. Quantide specializes in IT services for statistical purposes. The company is a leading expert in "R".

Work experience

Lecturer of “Quantitative Methods” course for MIB - Master in International Business

| | |
|-----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Period | August-September 2020 |
| Client | Università Cattolica del Sacro Cuore |
| Role | Lecturer |
| Main activities | <p>“Quantitative Methods” is a fully online pre-course that is preliminary to “MIB Regular and Executive”.</p> <p><i>The purpose of this short course is to provide students with key concepts of quantitative methods, with the aim of making them familiar with the main concepts and basic introduction to descriptive and inferential statistics, highlighting the difference between correlation and causation, and introducing probability and hypothesis testing theory.</i></p> <p>Development of 8 video lessons for the pre-courses of Academic Year 2020 (held completely online). Output of the course was 8 video lessons, slides, final test with solutions, suggested books and readings.</p> |

Lecturer of R for Data Science for PhD Students and Executives

| | |
|-----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Period | July 2020 |
| Client | Milano Bicocca University |
| Role | Lecturer |
| Main activities | Frontal lessons, including content planning, material preparation, written and oral exam . The course presented advanced topics to write algorithms and code in R in an effective way for Data Science purposes . Topics were: Tidyverse, functional programming, S3 object oriented programming, packages, parallel computing and performance optimization. Collaborative Programming with Git and Github. |

Organization of eRum 2020

| | |
|-------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Period | 2019-2020 (preparation period) |
| Event | 2020 European R Users Meeting, on-site conference turned into a virtual event: a 4 day international R conference with 2.000 attendees and ~150 speakers. See http://2020.erum.io/ |
| Role | Organizer, “R in Production” Track coordinator and Tech coordinator |
| Description | As an enthusiast of statistics and R technology, I have followed the international community for years. To organize the conference we founded the association Milano R. We almost completely organized a live conference until the sad appearance of the covid-19 a very short time before the event. This forced us to turn the event online. However, it was a success with 2000 attendees from all over the world and world-renowned speakers. |

Full stack Data Scientist

| | |
|--------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Period | June 2019 to Present |
| Role | Data Scientist and IT for data Science. |
| Main activities and responsibilities | <p>In this role I provide automated systems that access or store data, do some analysis with different pieces of software cooperating, and provide reports or dashboards.</p> <p>Examples are 2 projects. Scala API development for an Italian Banking institution .Scala Application provides real time manipulation and aggregation with great effectiveness and speed data provided by API.</p> |

And a second example for a **US Financial company** I provided a **Full Stack Reporting system**. I first developed some web reports, secondly I provided a customized web and R server which provides reports to authenticated clients. Finally I'm still developing other reports and updating them by continuous integration and delivery process.

Lecturer of Machine Learning & Big Data for Master in Data Science, Turin University

Period **May 2019**

Role Lecturer

Main activities and responsibilities Frontal lessons, including content planning and material preparation. The class was an introduction to big data and machine learning methods for big data, and to Big Data Software as well as R libraries aimed at this scope.

Full stack Data Scientist

Period **June 2018 to 2019**

Role Data Scientist and IT for data Science.

Main activities and responsibilities In this role I provide automated systems that access or store data, do some analysis with different software cooperating, and provide reports or dashboards. Examples are 2 projects for "**Energy Team**". The first is an **RMarkdown automatic HTML report** sent by email. This markdown is a monthly report scheduled to be compiled automatically and sent every first of the month to a number of clients. It allows users to read data interactively (plotly and other js libraries) and extract some Excel summary from it. This has a very nice aesthetics and a thorough attention to details such as the unit of measures, scales and colour usage. The second is a **R Report generator API** integrated with a Java Web platform. The report allowed the interaction between engineers and Data Scientists through the Platform. Java was the interface and R with R Plumber the complete graphics and report generation engine. Data Scientist and Data Visualization Consultant. A Project example with **Revelo DataLabs** where I improved their dashboard with a better visualization and automated an analysis in the background of the Dashboard.

Information Technology and Services Consultant for Quantide S.r.l.

Period **June 2015 to 2019**

Role IT Management Consultant & R Trainer

Main activities and responsibilities In this role I have responsibility for the functionality of the production environment and of its recovery systems. I also provide installation and configuration of systems with R, Hadoop, Spark, Hive, and Sparklyr used on the Cloudera platform. IT systems are automated with configuration scripts: they automatically set up each server responding to its role in the infrastructure. I also code "R", Shiny and train data analysts how to code "R" efficiently. As a Shiny professional I completed assignments utilising "Shiny" for Global multinationals such as **Microsoft, UniCredit Bank, Lavazza Eli Lilly, Enel Energy, TetraPak, and Radici Group**. My primary responsibility is an ongoing consultancy service for Altroconsumo, a major Consumer Service Group. This has been an ongoing service provided by Quantide, which started six years ago. I have been working on this assignment for three years with a time commitment of two/three days per week. Also, as the co-founder of R-Labs I spearheaded two projects, Comune di Milano, and Earth Cloud, wherein Shiny was used to provide valuable solutions in each case.

Lecturer of R classes for Companies and Open Courses

| | |
|--------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Period | June 2015 to 2019 |
| Role | Lecturer |
| Main activities and responsibilities | <p>Lessons include materials and exercise sessions for a group of 6 to 8 attendees.</p> <ul style="list-style-type: none"> - R for Data Science: first R session to data manipulation, visualization and discovery, from data import and export to data modelling and data mining: a wide overview of R as a data science tool. - Data Visualization and Dashboard with R (I developed the material): data visualization techniques with R and how to export results through reports and dashboards. - R with Database and Big Data (I developed the material): manipulating, analyzing, visualizing big data structures that exceed the single computer capacity in a true R style. - Professional Programming with R: how to become a R developer with a deep knowledge of R and the modern techniques of programming. <p>Lecturer of “Advanced Programming for Big Data” at MABDA (Master in Big Data Management) in LUISS University Rome</p> |
| Period | September 2016 |
| Role | Lecturer |
| Main activities and responsibilities | I taught about how to analyze big data with Hadoop, Apache Spark and R. |
| | Lecturer of “R for Large Computation” for Master Bocconi University Milan (Research Division) |
| Period | 2/3 November 2015 |
| Role | Lecturer |
| Main activities and responsibilities | I taught a one day class for the introduction to linux working toolbox (Frontal lessons, including content planning and material preparation). Then I worked as professor assistant for the second day. |
| | Information Technology and Services for ESIT S.r.l. |
| Period | 2006 to 2007 |
| Company | ESIT S.r.l |
| Main activities | Hardware and software support services for several major companies including financial services organisations, Banking institutions, and industrial organisations. |
| | Information Technology and Services for "Università degli studi di Milano". |
| Period | 2008 to 2009 |
| Company | "Università degli studi di Milano" |
| Main activities | Multimedia calculus lab of "Università degli studi di Milano" |
| | Master Degree in Physics -- Università degli Studi di Milano |

Graduation **2016**
 Thesis title "Universality and specificity in human communication dynamics. The case of reddit."
Bachelor Degree in Physics -- Università degli Studi di Milano
 Graduation **2011**
 Thesis title "Procedure di reweighting nella stima dell'incertezza dovuta alle distribuzioni partoniche del protone"

Information Technology Skills

R Programming **Professional and advanced programming** with statistics, big-data (Hadoop-Spark) scope. **Shiny** interfaces and dashboards. **Report automation. Integration of C/C++,Java, Scala** with R. **Training.**

Cluster Hadoop Management of a Hadoop cluster with and without **Cloudera**. Configuration of Spark and R for Data analysis.

Big Data analysis with Spark Installation and configuration of **Spark upon Hadoop** with **Sparklyr** as R interface or **Scala API**.

Server Configuration of **Linux R server** (ssh, ftp, vnc, proxy, vpn, **RStudio-server, RStudio-Shiny-Server-Professional**). The server has statistical purposes and allows collaboration between several users using R and other statistical tools.

Amazon Web Services Creation, configuration, orchestration of group of virtual machines with **AWS_CLI**, Python Boto3, Vagrant. Solid experience with AWS Services: EC2, VPC, S3, Elastic Map Reduce, EBS, CodeCommit

DevOps Skills Infrastructure as code with **Ansible, Puppet, Vagrant**. Automation and orchestration with Puppet.

Networks Knowledge of principles of networking, firewalling. Hardware and software configuration of routers, switches and Wireless Access Points.

Google Web Services Apps-for-work and Google Cloud (especially **BigQuery**)

Cluster Debian Linux Management of 400 actual users, 50 client terminals, 1 proxy (firewall, and server html), 1 NIS server with raid software and log and backup server.

C and C++ Programming Especially in scientific area: simulations or complex calculation. I provide also C/C++ integration with R.

Other programming languages **Scala, Python** primarily. And others like Haskell, Mathematica, Fortran, Maxima, Gnuplot, GraphViz (dot language)

Programming tools **Git, Github, Travis, Svn, Emacs, Vim**

Scripting **Bash** and Zsh (Advanced Scripting and scheduling)

Projects

Organizing the European R User Meeting 2020 in Milan

I am in the organization committee of the eRum2020 a 3 day international R conference with about 500 participants and ~100 contributors. See <http://2020.erum.io/>
 I participate in the Admin group (As a member of the MilanoR Association) and I am the responsible of the R-in-production Track of the program.

R-Lab Milano

The R-Labs are monthly group meetings where users meet to collaborate on the resolution of a given problem related to data science, which would have occurred in our territory and proposed by a company or university or organization. The goal of each meeting is to develop a feasible solution for the problem, and publish the developed code through Github.

The goal of the R-lab consists of three components:

- To strengthen the R community by facilitating the interaction and the exchange of knowledge between users.
- To pursue the broadening of knowledge of our users, offering real problems stemming from a wide range of fields and experiences.

To push the open-source and R culture in our area, showing how R, open source and the community are useful resources for solving problems. We want to emphasize the concept that through collaborative processes it is possible to discover more solutions and quicker.

Achievements

Spearheaded two projects with R Labs using Shiny to achieve the following:

1. Comune di Milano: Provided a solution to facilitate a method for distributing financial information to the public.
2. Earth Cloud: Provided a solution to analyse Seismic activity and to create an early morning Earthquake Alarm System.

Personal Interests and Activities

Karate

More than 15 years of practice of traditional karate. National meetings and competitions. Teaching of Karate discipline at evening classes and at primary schools. Active participation in the organizing of the sport's association. I have been Vice-president of Bu Do Kan Karate in Samarate and Vice-President of Asso-Sport (non-profit organization that organizes and coordinates several sport organizations of Samarate)

Personal skills and competences

Mother tongue(s)

Italian

Other language(s)

Self-assessment
Understanding
European level

| | Understanding | | Speaking | | Writing |
|----------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | Listening | Reading | Spoken interaction | Spoken production | |
| English | B2: Independent user | B2: Independent user | B2: Independent user | B2: Independent user | B2: Independent user |
| French | B1: Independent user | B1: Independent user | B1: Independent user | B1: Independent user | B1: Independent user |

(*) *Common European Framework of Reference for Languages*