



(Italian version below)

**Proposal for Improving the Educational Offering in the *PhD Program in Psychological Sciences*,
*University of Padua***

Free Software and Python

February 03, 2024

In view of the upcoming deadline for approving the teaching plan for next year's PhD program, and the forthcoming **CFU/ECTS system** for courses, which I expect will increase the willingness of faculty members to offer courses, I would like to propose a measure to improve the educational offering in the PhD program in Psychological Sciences. This proposal stems from a series of recent discussions with colleagues, faculty members, and doctoral students. I am available to help directly with its implementation, as far as possible.

The proposal has several objectives: 1. To **improve the coherence** of the current educational offering; 2. To support a more **efficient use of financial resources**; 3. To enhance the **versatility of the technical training** of future researchers. It also stems from the observation that there is now a wide range of excellent alternatives (e.g., R, JASP, Jamovi, Python) to costly proprietary software (e.g., SPSS, E-Prime).

The proposal is as follows:

- Standardize the teaching of **mandatory foundational courses** in our PhD School **to ensure that they rely exclusively on free and open-source software**;
- Introduce a basic course of **Python**;

Additionally, we suggest that:

- The use of **proprietary licensed software** in PhD courses be **restricted to optional courses** and must be **explicitly justified** by the absence of viable alternatives.

The first suggestion is motivated by the consideration that, **given the current availability of free software, it is inappropriate to require doctoral students to train using tools they may not easily access in their future careers**. Courses involving proprietary software can be offered as *optional*, but we believe their use should be explicitly justified. In particular, it is advisable to avoid using costly licensed software to teach skills that are **independent of any specific tool** (e.g., statistical analysis, general programming skills).

It has also been informally reported by some doctoral students that the use of different software for the same tasks across courses **can lead to confusion**, especially when the courses are thematically related or sequential (e.g., on statistical analyses). If a choice must be made, it is clearly preferable to adopt a consistent approach based on free and open-source languages. R is the gold standard for *statistical analysis*; Python is the equivalent for broader *programming* and *data science* tasks.

The second suggestion (Python) stems from the observation that the PhD program should prepare future researchers to work not only in academia but also in professional and industry contexts. For this reason, it is important to at least introduce students to the most widely used software in data-driven industries. For years, **Python has been the most commonly used language for data science worldwide**, especially in **corporate settings**, and it is **free**.

With the growing use of **artificial intelligence** and **machine learning**, we may expect that having at least some basic knowledge of **Python** will be a **fundamental asset** for future researchers, particularly for those who pursue careers outside academia.

Kind regards,
Enrico Toffalini

Proposta di miglioramento dell'offerta formativa nel dottorato in Psychological Sciences

Software libero e Python

03/02/2024

Alla luce della ravvicinata scadenza per l'approvazione della didattica del dottorato del prossimo anno, e della futura **creditizzazione degli insegnamenti**, che presumo contribuirà ad ampliare la disponibilità dei docenti a svolgere corsi, vorrei avanzare una proposta per il miglioramento dell'offerta formativa in *Psychological Sciences*. La proposta nasce da una serie di recenti confronti con colleghe/i, docenti, e anche dottorande/i. Mi metto a disposizione per contribuire a una realizzazione diretta di questa proposta, per quanto possibile.

La proposta ha diverse finalità: 1. **migliorare la coerenza** dell'attuale offerta formativa; 2. favorire il **risparmio di risorse economiche**; 3. massimizzare la **versatilità della preparazione tecnica** delle future ricercatrici e ricercatori. Nasce inoltre dalla constatazione che oggi esiste un'ampia offerta di ottime alternative (es. R, JASP, Jamovi, Python) ai software proprietari con licenza onerosa (es. SPSS, E-Prime).

La proposta consiste in questo:

- Uniformare l'offerta didattica dei **corsi base obbligatori** della Scuola di dottorato in modo tale che prevedano esclusivamente l'**uso di software liberi e gratuiti**;
- Includere un **corso base di Python**;

Inoltre suggeriamo che:

- **Software con licenza onerosa** nei corsi del dottorato siano limitati ai **corsi opzionali**, e vadano esplicitamente **giustificati** dall'assenza di alternative;

Il primo suggerimento è motivato dalla riflessione che, **alla luce dell'attuale offerta di software liberi, non è opportuno vincolare dottorande/i a una formazione che impiega software di cui potrebbero non disporre facilmente nel loro futuro lavorativo**. Eventuali corsi per l'uso di software proprietari possono essere offerti come *opzionali*, ma crediamo andrebbero esplicitamente motivati. Crediamo infatti sia meglio evitare l'uso di software con licenza onerosa per insegnare competenze che prescindano dal loro **specifico uso** (es. competenze statistiche o di programmazione generale). Mi è stato inoltre riportato informalmente da alcune/i dottorande/i che l'uso di software diversi per diversi corsi tende a **creare confusione**, soprattutto laddove i temi siano collegati e propedeutici l'uno all'altro. Dovendo scegliere, è chiaramente preferibile uniformarsi sull'uso di software/linguaggi liberi. R è il gold standard per l'*analisi statistica*, Python lo è per tutte le altre esigenze di *programmazione e data science*.

Il secondo suggerimento (Python) nasce dalla constatazione che, dovendo il dottorato formare future ricercatrici e ricercatori capaci di applicarsi non solo nell'ambito accademico, ma

anche in quello professionale e aziendale, è importante almeno avviare alla conoscenza del software più usato nelle realtà aziendali che elaborano dati. Da anni Python è il **software globalmente più diffuso** per *data science* specialmente nelle **realtà aziendali**, ed è gratuito. Con il diffondersi dell'uso di **intelligenza artificiale** e *machine learning*, crediamo che possedere almeno una base di **Python** sia un **asset fondamentale** per le future ricercatrici e ricercatori, soprattutto per coloro che si applicheranno al di fuori dell'ambito accademico.

Cari saluti,
Enrico Toffalini