Italo M. Duran

Harrison, NJ 07029 | 786-521-8403 | italo.md993@gmail.com| [LinkedIn](https://www.linkedin.com/in/italoduran/) | Github: [imd9](https://github.com/imd9)

# work history

**Johnson & Johnson** – Ethicon Raritan, NJ

RDLDP Engineer Data Science / Bioinformatician *02/2023-11/2024*

* **Developed 'Janytor'**, a Python Dash application for automated data cleaning & transformation, reducing data processing time from 30 minutes to **7 seconds** per dataset.
  + Eliminated a 24-hour delay in data availability, enabling immediate post-experiment analysis.
  + Achieved annual time savings of over **3,600 hours**, equivalent to **$183,000** in cost savings.
* Developed an **R Shiny app** for data transformation, summary statistics, and data visualization.
* Maintained and updated data analysis platforms, enhancing operational efficiency.
* Assisted in developing Python scripts for the NPDU Python course.

# Projects & Research experience

**Biological Database web development for Mutational Accumulation of yeast**

• Aiming to develop an open-source web interface that relocates and pools raw data files, cross references feature and performs

statistical aggregation to easily visualize queries against experimental data and automate workflows

• Worked in python and SQL to build a database focusing on identifying C to A mutational frequency in yeast strains containing msh2 mutations

**Extended analysis of Dengue Virus-2 genome**

* Assessed an extended analysis of Dengue Virus-2 genome by structure, domains, and phylogeny analyses
* Built an array package with several methods and attributes in python to perform calculations

**Microarray based tumor classification**

* Reproduced a colorectal cancer microarray analysis, focused on differentiating C3 and C4 tumor subtypes
* Developed a pipeline to process microarray data including normalization, batch correction, and scaling
* Employed Principal Component Analysis to reduce dimensionality and examine outliers using ggplot2

**Analysis of transcriptional profile of murine cardiac regeneration using mRNA-Seq data**

* Performed quality control on FASTQ files and recreated figures from O’Meara et. al., 2015
* Obtained mRNA-Seq datasets from GEO to generate estimates of transcript abundance/depletion in postnatal day 0 (P0) and adult (Ad) murine cardiomyocytes utilizing CuffLinks package
* Compared FPKM counts of genes involved in three biologically distinct functional processes using hierarchical clustering algorithm and heatmaps to interpret upregulation / downregulation between datasets

**Analysis of neuraminidase and nucleoproteins of H1N1 influenza type A**

* Conducted a multiple sequence alignment study to inspect and predict active target sites on neuraminidase and nucleoproteins of H1N1 influenza type A

# Education

Boston University Boston, MA

MS, Bioinformatics 2021-2023

Florida International University Miami, FL

BS, Biology 2017 – 2019

# skills / certifications

* **Programming:** Python, Html, CSS, SQL, mySQL, R, Rshiny, Excel, Dash, plotly
* **UX/UI:** HTML, Figma, CSS, Adobe (lightroom, Photoshop, InDesign)
* **Databases:** PDB, UniProt, Clustal Omega, SCOP, STRING, KEGG, GEO, GenBank
* Working experience with TCGA, protein, NCBI, Ensembl, EBI databases
* **Machine Learning**: K-means, K-Nearest Neighbor, Dimensionality reduction algorithm, PCAs
* Excellent communications and organizational skills across multiple fields.
* **Languages:** Fluent in English & Spanish; Elementary Japanese