

PROGRAMA DE CURSO

Código	Nombre			
IN7G4	Ciencia de datos para mejores políticas públicas			
Nombre en Inglés				
Data science for better public policies				
SCT	Unidades Docentes	Horas de Cátedra Semanales	Horas Docencia Clase Auxiliar Semanales	Horas de Trabajo personal Semanal
3	5	3		2
Requisitos			Carácter del curso	
			Electivo MGPP	
Resultados de Aprendizaje				
<p><i>Al término del curso el/la estudiante demuestra que:</i></p> <ul style="list-style-type: none"> - Comprender y aplicar nociones de matemática, estadística y programación, al servicio de la Ciencia de Datos. - Comprender y aplicar buenas prácticas en el diseño de modelos según los objetivos y preguntas planteadas. - Comprender y aplicar críticamente diversas técnicas de análisis de datos, reconociendo su pertinencia en función de los objetivos de análisis y datos utilizados. - Comprender y aplicar análisis geoespaciales básicos, al servicio de la construcción de mejores modelos. 				
Metodología Docente			Evaluación General	

UNIDADES TEMÁTICAS

Número	Nombre de la Unidad	Duración en Semanas
1	Introducción a la Ciencia de Datos	0,5
Contenidos	Resultados de aprendizaje de la Unidad	Referencias bibliográficas
Introducción a Ciencia de Datos, <i>Métodos de definición de objetivos</i> y Python.	<p><i>El/la estudiante:</i></p> <ul style="list-style-type: none"> - Conoce conceptos y definiciones de Ciencia de Datos, objetivos de diseño y una introducción a Python. - Declara los conceptos y definiciones a utilizar a lo largo del curso, dando particular importancia al diseño desde su origen. - Se introduce en la principal herramienta a utilizar: Python. 	
Número	Nombre de la Unidad	Duración en Semanas
2	Adquisición, exploración, preparación y minería de datos Python	2,5

Contenidos	Resultados de aprendizaje de la Unidad	Referencias bibliográficas
<p>Adquisición y preprocesamiento de datos asegurando una mínima calidad de la información a trabajar, utilizando Python.</p> <p>Preparación de datos con Python: estructura y filtros.</p> <p>Minería de datos con Python.</p>	<p><i>El/la estudiante:</i></p> <ul style="list-style-type: none"> - Comprende la correcta elección de entidades, entendimiento de sus atributos y el diseño de tratamientos para los datos, previo a cualquier análisis de la herramienta Python. - Maneja la limpieza y preparación de la información en su formato y forma según objetivos del diseño, asegurando la relevancia del insumo previo a análisis y procesos. - Aplica métodos estadísticos u otros métodos avanzados para identificar patrones, utilizando Python o R. 	

Número	Nombre de la Unidad	Duración en Semanas
3	Modelación de datos para Políticas Públicas	3
Contenidos	Resultados de aprendizaje de la Unidad	Referencias bibliográficas
<p>Planificación y diseño de modelos.</p> <p>Desarrollo y validación de modelos.</p>	<p><i>El/la estudiante:</i></p> <ul style="list-style-type: none"> - Es capaz de definir el diseño de modelos de acuerdo con el objetivo de diseño declarado y consolidado, involucrando una vista global del ciclo de Ciencia de Datos, y manteniendo siempre en el radar cuál es el objetivo. - Desarrolla modelos, de acuerdo con el objetivo, aplicando ML y/o HPC utilizando Python o R. - Maneja herramientas de validación de modelos en sus puntos críticos. 	

Número	Nombre de la Unidad	Duración en Semanas
3	Evaluación, rediseño, interpretación y comunicación de resultados	2

Contenidos	Resultados de aprendizaje de la Unidad	Referencias bibliográficas
<p>Evaluación y rediseño.</p> <p>Interpretación y comunicación de resultados a través de visualizaciones.</p>	<p><i>El/la estudiante:</i></p> <ul style="list-style-type: none"> - Analiza la calidad y detecta falencias en modelos desarrollados. - Es capaz de generar adecuaciones al proceso de análisis del modelo. - Interpreta correctamente los resultados y/o hallazgos. - Es capaz de comunicar los resultados a través de herramientas y métodos de visualización de datos interactiva. 	

Bibliografía General

Obligatoria:

- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Edelman, B. G., & Luca, M. (2014). Digital Discrimination: The Case of Airbnb.com (SSRN Scholarly Paper No. ID 2377353). Rochester, NY: Social Science Research Network. Disponible en <https://papers.ssrn.com/abstract=2377353>
- Goldsmith, S., Crawford, S., & Weinryb Grohsgal, B. (2016). Innovaciones en la prestación de servicios públicos: Número 4: El análisis predictivo: Impulsando la mejora a partir de los datos (Discussion Papers & Presentations). Inter-American Development Bank. <https://doi.org/10.18235/0000421>
- Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. *The American Economic Review*, 105(5), 491–495. <https://doi.org/10.1257/aer.p20151023>
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big Data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute. Disponible en <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation>
- Rodríguez, P., Palomino, N., & Mondaca, J. (2017). El uso de datos masivos y sus técnicas analíticas para el diseño e implementación de políticas públicas en Latinoamérica y el Caribe (Discussion Papers & Presentations). Inter-American Development Bank. <https://doi.org/10.18235/0000694>
- Tomar, L., Guicheney, W., Kyarisiima, H., & Zimani, T. (2016). Big Data in the public sector: Selected applications and lessons learned. Inter-American Development Bank. Disponible en <https://publications.iadb.org/handle/11319/7884>

Complementaria:

- Anderson, C. (2008, junio 23). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. Disponible en <https://www.wired.com/2008/06/pb-theory/>
- Bahamonde, J., Hevia, A., Font, G., Bustos-Jiménez, J., & Montero, C. (2014). Mining Private Information from Public Data: The Transantiago Case. IEEE Pervasive Computing, 13(2), 37–43. <https://doi.org/10.1109/MPRV.2014.30>
- Butler, D. (2013). When Google got flu wrong. Nature News, 494(7436), 155. <https://doi.org/10.1038/494155a>
- Friedman, V. (2008, enero 14). Data Visualization and Infographics – Smashing Magazine. Disponible en <https://www.smashingmagazine.com/2008/01/monday-inspiration-data-visualization-and-infographics/>
- Harford, T. (2014, marzo 28). Big data: are we making a big mistake?. Disponible en <https://www.ft.com/content/21a6e7d8-b479-11e3-a09a-00144feabdc0>
- Hilbert, M., & López, P. (2011). The world's technological capacity to store, communicate, and compute information. Science, 332(6025), 60–65. <https://doi.org/10.1126/science.1200970>
- The Economist. (2010, febrero). Data, data everywhere | The Economist. Recuperado de <http://www.economist.com/node/15557443>
- Miller, C. (2015, julio 9). When Algorithms Discriminate. The New York Times. Disponible en <https://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html>

Recursos Web

- J. Vanderplas, The Jupyter Notebooks behind my O'Reilly report, "A Whirlwind Tour of Python": [jakevdp/WhirlwindTourOfPython](https://github.com/jakevdp/WhirlwindTourOfPython). 2018.
- R. Kumar, IPython Notebooks to learn Python. Contribute to [rajathkmp/Python-Lectures](https://github.com/rajathkmp/Python-Lectures) development by creating an account on GitHub. 2018.
- "Jupyter Notebook Viewer". [Online]. Disponible en: <https://bit.ly/2CqpH9h> .
- "Python for Developers". [Online]. Disponible en: <https://bit.ly/2oNFs0G> .
- "Tutorial de Python 3.6.3". [Online]. Disponible en: <https://bit.ly/1IrfUbn>.