

ENTREVISTA A ÓSCAR RUEDA

Óscar Rueda es un matemático y estadístico español que trabaja en la Universidad de Cambridge para la investigación contra el cáncer de mama. Su rama de especialización es la bioestadística.

PREGUNTAS:

1. ¿Cuál es su formación académica? ¿Qué le ha conducido a Cambridge?

Soy Diplomado y Licenciado en Estadística por la Universidad de Valladolid. Hice el doctorado en el CNIO en Matemáticas (rama Estadística). Después me fui de investigador postdoctoral a CRUK en Cambridge, al grupo de Carlos Caldas en Genómica Funcional del Cáncer de Mama.

2. ¿A lo largo de sus años de formación académica tuvo usted la oportunidad de aprender a programar en R o, por el contrario, ha sido más bien un aprendizaje independiente y complementario?

Sí, tuvimos una asignatura en la carrera de Estadística en la que dábamos R, allá por el año 2000. Después de haber utilizado un montón de paquetes estadísticos y programar en varios lenguajes, descubrir R fue un gran cambio, ya que parecía estar diseñado desde el punto de vista del estadístico.

3. ¿Ha empleado usted R en algún otro trabajo aparte del actual?

Sí, lo he utilizado constantemente durante los últimos veinte años.

4. ¿Qué ventajas ofrece R frente a otros lenguajes de programación como Python o Matlab?

Bueno, yo creo que la elección de un lenguaje frente a otro viene determinada por gustos personales, y sobre todo a lo que estamos más acostumbrados. Para mí, las principales ventajas de R son que está diseñado para el estadístico, que tiene una infinidad de paquetes con todas las técnicas que puedo necesitar, que permite modificar o personalizar cómodamente cualquier función y que me permite poner a disposición de la comunidad cualquier método que produzca de manera fácil y rápida.

5. Si pudiera cambiar cualquier cosa que no le agrada especialmente de R, ¿qué cambiaría?

R tiene problemas con data sets demasiado grandes, pero está mejorando en eso también.

6. ¿Cree usted que R es un lenguaje de programación en auge? ¿Cuál cree que será su importancia dentro de 5 años?

R sigue estando muy de moda, aunque noto cierta tendencia a favor de Python. También veo que las nuevas generaciones usan demasiado los nuevos entornos (tidy verse, ggplot, etc) que a mi modo de ver se alejan demasiado de la idea original de R, aunque es una forma de hacerlo más atractivo para usuarios que vienen de otros entornos menos estadísticos.

7. ¿Qué importancia cree usted que tiene el saber manejar R para alguien que trabaja en la investigación?

Mucha, aunque trabajes con Python o Matlab, es muy probable que tengas que utilizar algún método que está implementado en R. Además, para hacer análisis rápidos y producir gráficos de alta calidad es una herramienta muy útil.

8. ¿Qué fue lo que despertó su interés en la estadística, y en concreto en la bioestadística orientada al estudio genómico?

Fue un curso de Estadística que di mientras estudiaba Empresariales. Me gusto tanto la asignatura, en parte gracias a su profesor que era estupendo, que decidí empezar la carrera de Estadística. Durante los siguientes años, tenía claro que quería dedicarme a aplicaciones relacionadas con Medicina, y durante mi estancia en el CNIO por supuesto descubrí la Genómica y me pareció un campo apasionante, con multitud de problemas interesantes para un Estadístico.

9. En su vida cotidiana, ¿tiene momentos donde utilice R para fines no estadísticos?

Pues no se me ocurre ninguno hasta ahora.

10. ¿Qué papel juega R en su día a día como estadístico que estudia el genoma humano?

¡Prácticamente todos mis análisis se hacen en R!

11. Finalmente, ¿podría hablarnos un poco de su investigación contra el cáncer de mama? ¿En qué consiste el proyecto que lleva a cabo en el CRUK?

Bueno, ya no estoy en el CRUK, ahora tengo mi propio grupo en la Universidad de Cambridge dentro del MRC-BSU. Allí me dedico a diseñar métodos para detectar biomarcadores que permitan estratificar pacientes de cáncer de mama, de forma que se puedan identificar tratamientos específicos para ellos. También hacemos modelos para monitorizar el estado de pacientes y predecir el riesgo de metástasis utilizando información genómica del tumor. En general, queremos encontrar las características que hacen que cada tumor sea único y utilizar esas características para predecir qué tratamiento va a ser más beneficioso.