

# Sports Analytics

Incentivando la actividad física usando DATA

# Raimundo Sánchez, PhD.

- Ingeniero Industrial, Doctor en Sistemas Complejos.
- Profesor de Data Science y Advanced Analytics en UAI.
- Investigador en Sports Analytics.
- 10 años de experiencia en el desarrollo de herramientas basadas en datos para resolver problemas del mundo real.
- Entre 2015 y 2019, Gerente de Revenue Management Analytics en LATAM Airlines, a cargo de 13 ingenieros.
- Corredor aficionado:
  - 3 x 50km TNF EC Chile,
  - 2 x 42km Maratón de Santiago (PB 3:10:00)

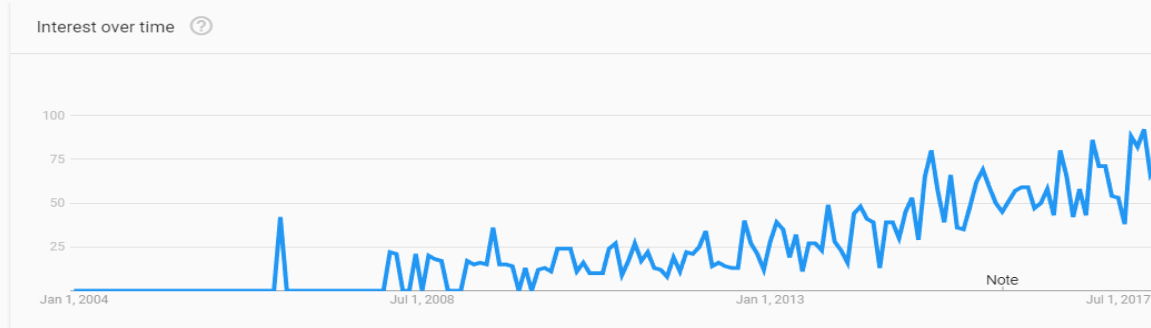


# Actividad física en Chile.

- El comportamiento sedentario es la tercera causa de muerte en el mundo, después de las enfermedades cardíacas y el cáncer.
- El 81% de las personas en Chile están físicamente inactivas.
- Un 10% de los chilenos practican el running / mountain running.

# Sports Analytics

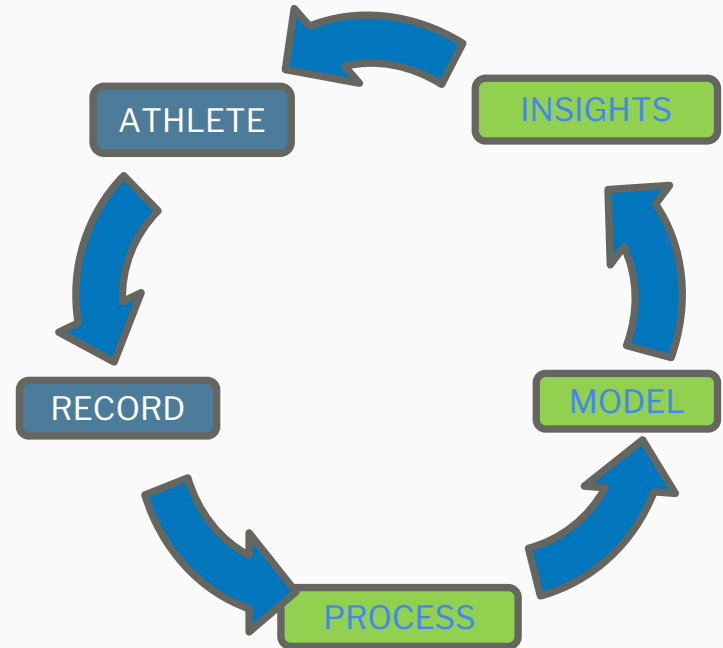
- En los últimos 10 años la disciplina se ha ido desarrollando
- Centrada principalmente en predicción de resultados de eventos deportivos
- Foco en atletas elite en competencias profesionales
- Existen diversas oportunidades para extender conocimientos a población general



# Retroalimentación digital

- La información generada por cada atleta puede ser utilizada para responder diversas preguntas.
- Dichas preguntas se pueden responder de manera digital a través de modelos matemáticos.
- Los análisis buscan **motivar a los atletas** a seguir realizando actividad física

## Gamification



# Grupo Sports Analytics UAI

Estamos formando el equipo de Sports Analytics UAI, con foco en diferentes disciplinas deportivas:

- Deportes en equipo: Fútbol, Rugby, Basketball, Tenis, Golf
- Deportes individuales: Natacion, Trail running, Road running, Triatlon
- Deportes extremos, motorizados, etc.

Y generando acuerdos de investigación con diversas instituciones:

- CLC, PUC, CDUC, MinDep, etc.

# Charla Fútbol Analytics

Primer evento del Grupo Sports Analytics UAI

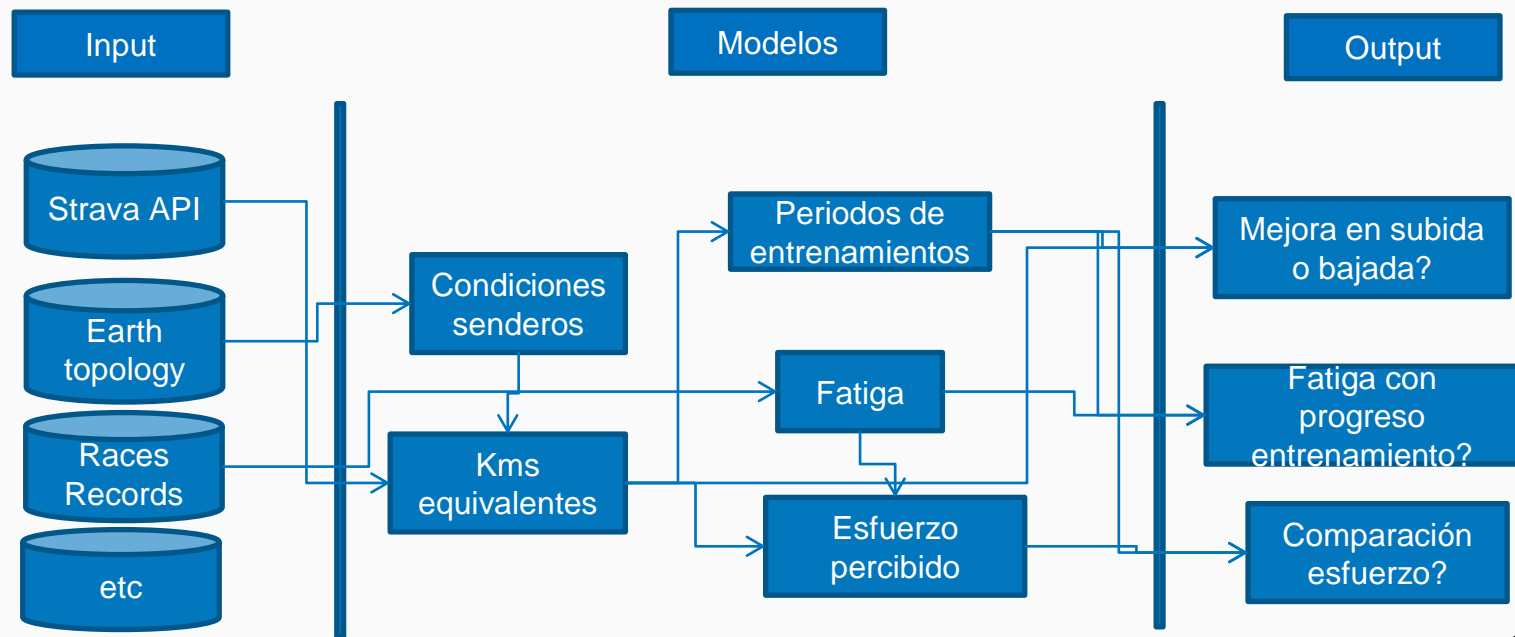
- Expositor: Ismael Gomez (@datoFutbol)
- Día antes de inauguración de copa américa
- 13 de Junio de 9:00 a 11:00
- Sede UAI Presidente Errazuriz

# Preguntas de Investigación

- ¿Qué métodos para capturar y procesar datos son necesarios para generar data confiable?
- ¿Cuál es la medida equivalente para la comparación de esfuerzos?
- ¿El conocimiento actual es representativo para corredores amateurs?
- ¿Cuál es la estrategia de entrenamiento óptima?
- ¿Cuál es la estrategia de competición óptima?

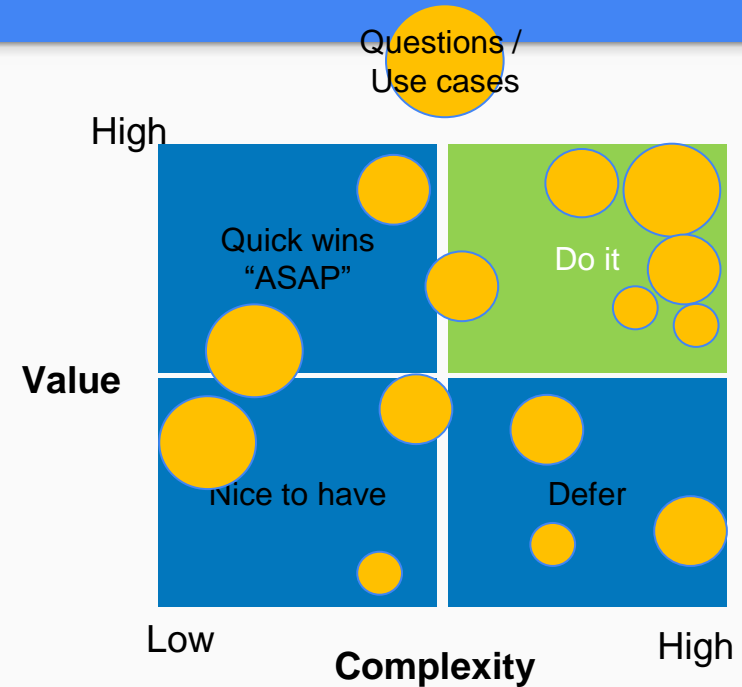


# Arquitectura de datos orientada a responder las preguntas planteadas



# Estrategia de datos

- Estamos estableciendo prioridades en torno a las preguntas que se quieren responder:
- Se realiza un mapeo de casos de uso según valor/complejidad.
- Priorizaremos el desarrollo equilibrado entre:
  - Valor académico
  - Valor aplicado



# Caso práctico: Mountain Running Analytics

# Corriendo en laderas

- El 64% del territorio de Chile corresponde a montañas.
- Número de carreras de montaña en Chile ha crecido 1000% en los últimos 8 años.
- La pendiente y el terreno de las rutas exigen una técnica diferente para correr que una superficie plana
- También cambia el gasto energético y la eficiencia metabólica.
- Una de las grandes preguntas en las carreras tiene que ver con estimar la duración de los eventos.



# Trail running analytics

Estudiar el rendimiento en carreras de montaña es extremadamente desafiante y requiere diversas fuentes de información.

- Existe una gran diversidad de terrenos y pendientes
- Diversidad de posibles rutas dentro de un terreno.
- Fenotipos y aptitud de atletas.
- Mecánica de suelos.
- Condiciones climáticas.
- Nutrición antes /durante carrera.
- Estrategias de alimentación y ritmo durante la carrera.

Todos estos factores son cuantificables

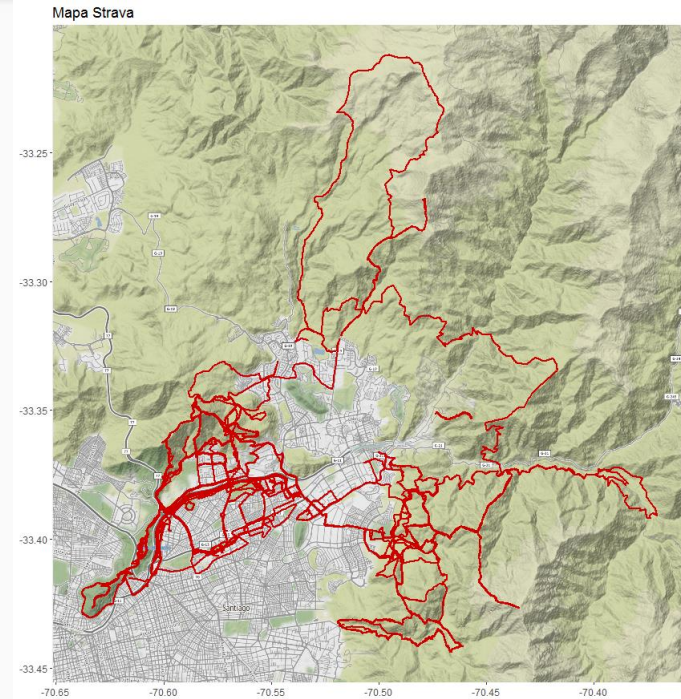
# Mountain Data

- A diferencia del atletismo tradicional, cada kilómetro en un evento de carrera al aire libre es diferente
- Para estudiar el rendimiento, no solo necesitamos el tiempo de finalización de cada evento, sino también la topología.
- La mayoría de la literatura en análisis de montañas usa estadísticas agregadas de organizadores de carreras
- La masificación reciente de dispositivos portátiles ha generado grandes cantidades de datos de actividad física
- La limitación de mayoría de estudios es que se enfocan solo en atletas de élite.



# Runners Data

- Los registros históricos de las carreras no tienen en cuenta los niveles de condición física de los atletas
- La mayoría de los autores consideran solo atletas de élite, controlando de esta manera el factor de entrenamiento.
- Algunos autores logran separar los datos de los diferentes niveles de condición física y logran desarrollar algunos modelos para todo el espectro de corredores.
- Las aplicaciones para correr permiten acceder a información histórica sobre entrenamiento de atletas y carreras.
- Requiere un registro disciplinado



# Equivalencia de desnivel con distancia horizontal.

- El modelo de equivalencia entre terreno plano y empinado más utilizado, es la fórmula de Naismith:  $EqKm = Km_{Hz} + \alpha Vert$
- $\alpha$  se conoce como el número de Naismith, la ITRA utiliza un valor 10 para asignar puntos a sus carreras.
- La principal limitación es que solo se ha estudiado dicho parámetro para atletas élites.
- Se han propuesto varias fórmulas alternativas desde entonces

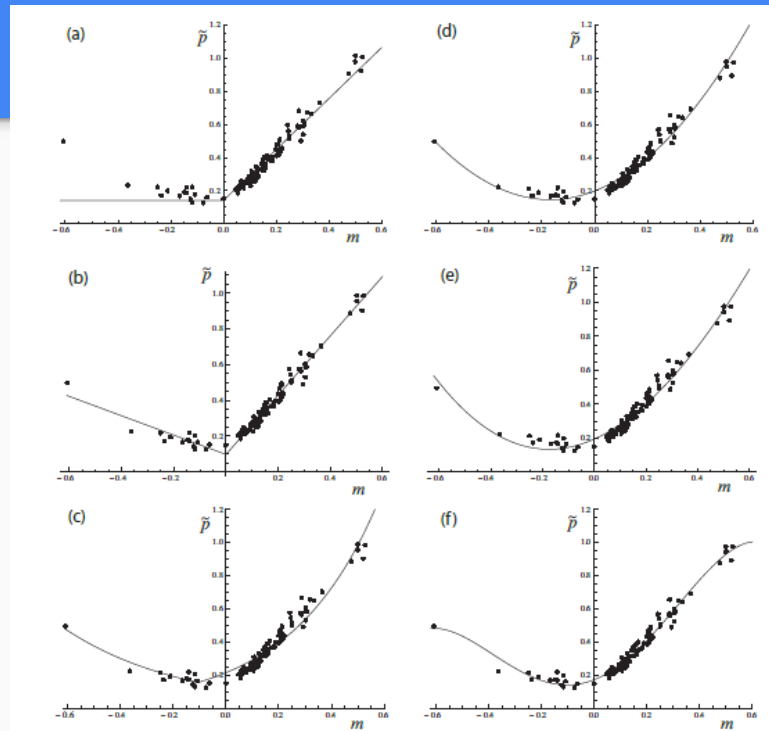
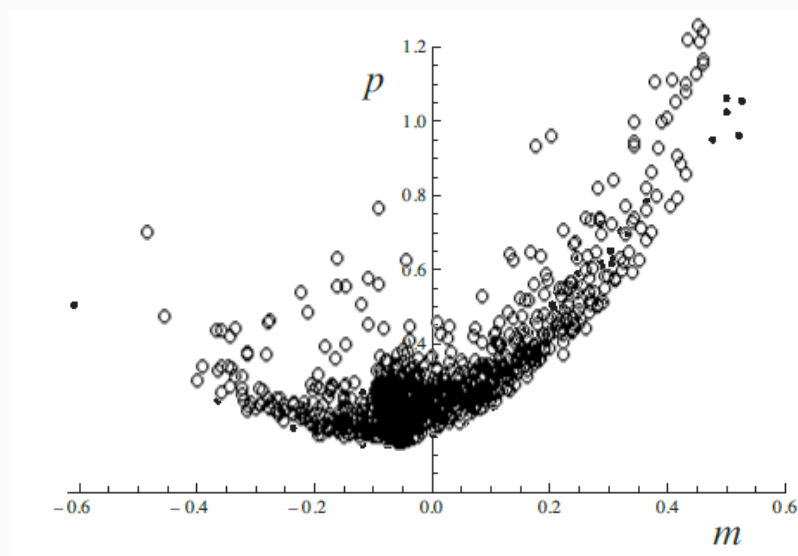


Figure 2: Adjusted pace data with various models fitted: (a) Naismith-type, (b) Piecewise linear, (c) Tobler-type, (d) Quadratic, (e) Cubic, (f) Quartic.



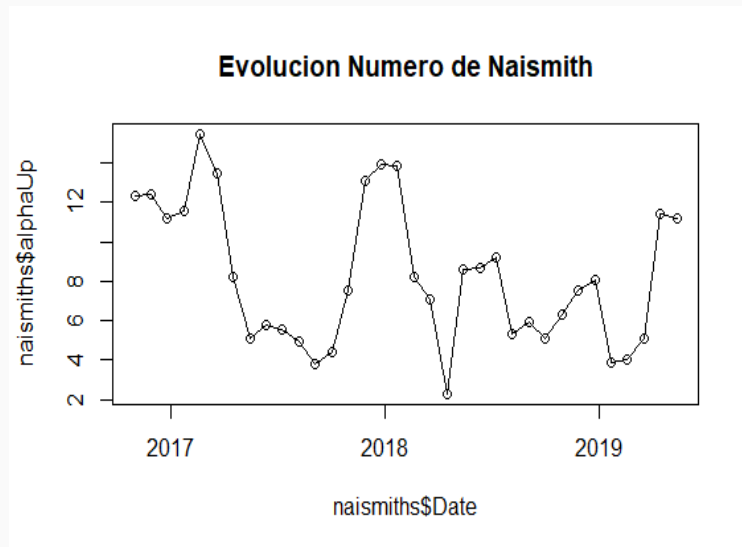
# Avances en equivalencia de kilómetros

- Gracias a la data GPS, se puede calibrar el modelo de Naismith para corredores de todo el espectro de desempeño.
- Programa se conecta a la API de Strava y extrae toda la data disponible para un usuario.
- Luego corrige la elevación de los tracks y elimina ruido de las trayectorias.
- Finalmente segmenta trayectorias y genera puntos de datos con mayor granularidad.



# Evolución Número de Naismith

- La variación del número de Naismith indica diferente evolución en el estado de forma de un deportista, siendo un  $\alpha$  más alto, *sinónimo de un mejor estado de forma*.
- Se implementó el calculo del número de Naismith para un periodo de entrenamiento móvil (12 semanas)
- Esto ayuda a medir la respuesta de los atletas a los entrenamientos.



# Próximos pasos

- Consolidación **red** de investigación en Sports Analytics
- Desarrollo de pruebas de concepto en diferentes disciplinas
- Postulación a diferentes fondos de investigación y transferencia tecnológica.

Hacer deporte!

# Sports Analytics

Incentivando la actividad física usando DATA