

PROYECTOS FINALES DEL SEMINARIO - 30%
CIENCIA DE DATOS EN EL MERCADO DE CAPITALES
12 de Octubre de 2018 - Facultad de Ciencias - UNAM

PUNTOS GENERALES

- Los equipos estarán conformados por 3 o 4 personas. Se asignará a cada equipo uno de los proyectos disponibles y tendrán una reunión con los instructores para definir las dudas principales. Para esto, por favor **manden a ambos instructores un correo** con propuestas del horario (**con al menos 3 días de anticipación**) para agendar una llamada en zoom.
- La fecha de entrega del proyecto es el **15 de diciembre (Sábado)**. Se deberá realizar una presentación del proyecto usando Zoom. **No habrá ningún tipo de prórroga.**
- Se tendrán que presentar dos avances, el primero el **30 y 31 de octubre** y el segundo el **20 y 21 de Noviembre**. **No habrá ningún tipo de prórroga.** Para presentar sus avances, cada equipo debe dar una exposición de 10 minutos durante el horario de clase. **Todos los integrantes** del equipo deben participar en estas exposiciones, la calificación será **por equipo y también de manera individual**.
- La información que se va a proporcionar para cada proyecto va a incluir los siguientes puntos:
 - Descripción del problema a resolver y resultado esperado
 - Temas involucrados y referencias
 - Dataset inicial. Se hará entrega de un dataset o el scraper básico, y el equipo deberá obtener el resto de las variables mencionadas por proyecto. En caso de incluir variables adicionales (alternative data) éstas deberán estar debidamente justificadas.

ENTREGA DEL PROYECTO

La entrega se realizará en un **repositorio de Github** por equipo (deberán enviar en **un sólo correo** todos los **nombres de usuario de Github** de los integrantes, a más tardar el día **19 de Octubre**). Estos repositorios van a estar dentro de la misma organización que las notas del curso:

<https://github.com/UNAMCiencias-Seminario-CienciaDeDatos>

Cada repositorio debe contar con:

1) **R markdown** describiendo la metodología (incluyendo fórmulas cuando sea necesario) y resultados. Este reporte debe incluir:

- Motivación y contexto teórico relacionado con el Mercado de Capitales
- Descripción de las etapas de trabajo, basado en el proceso general de minería de datos.
- Detalles de la extracción de datos adicionales (incluir la justificación).
- Detalles de la exploración de datos.
- Descripción de los modelos (supuestos, ventajas y desventajas).
- Discusión de los resultados en términos cuantitativos (al menos 3 gráficas).
- Discusión de los resultados en términos cualitativos: aplicado al Mercado de Capitales.
- Bibliografía, fuente de los datos y conclusiones.
- Incluir un resumen indicando claramente la colaboración que tuvo cada integrante.

- 2) Un folder con los **scripts**. El código debe estar comentado con procedimientos, desde la extracción de datos hasta la evaluación. El proyecto puede ser programado en R o Python.
- 3) Un folder con los **datos**: Formato csv o txt.
- 4) Una **presentación** de 10 min durante el horario de clase. La calificación de esta presentación será por equipo y también de manera individual.

DETALLES POR PROYECTO

1. Spotlight Index (Thematic Index)

Data: Asset Universe (Wikipedia S&P), Prices (quantmod), Analytics (Wikipedia)

Problem

Create a thematic index (Portfolio construction) using company page views of wikipedia (one single variable per company). The thematic index should have a basket of stocks which will be weighted considering the number of views a company has over the time.

The strategy must be done with a weekly-monthly rebalancing period and must include a backtesting from 2015.

Additional metrics must be included: VaR (Montecarlo, Parametric and Historical), Standard Shift (risk) and benchmark analysis (S&P).

Components: Web scraping, Backtesting, Portfolio Construction, Risk Metrics.

Expected Output: Thematic Index Strategy (holdings and weights) from 2015 to 2018, performance metrics, benchmark analysis, backtesting and risk metrics.

References:

- <https://www.mckinsey.com/industries/private-equity-and-principal-investors/our-insights/from-indexes-to-insights-the-rise-of-thematic-investing>
- <https://timtrice.github.io/backtesting-strategies/index.html>
- <https://pypi.org/project/pageviewapi/>