

WEB SCRAPING IN POWER BI

Indira Bandari



Microsoft Data Platform MVP



@indupriya9



in/indira-bandari-a1530119/



<http://mytechlearning.blogspot.co.nz>



indupriya9@yahoo.com

Indira Bandari

Experience

I have a Masters degree in Statistics and am passionate about data, analytics and learning data science.

Business Intelligence Consultant with over 15 years experience in designing and developing data warehouses and analytical solutions.

Speaker, Co-Organiser

Speaker at Local Users Groups and Meetups, webinars and SQL Saturdays in New Zealand. Co-Organiser for Auckland Artificial Intelligence Meetup, New Zealand Power BI Meetup, and Auckland SQL Saturday Event

Community

Volunteering and sharing my knowledge is my passion. In my pass time I teach game development for primary school kids. Last year I started teaching database concepts and data visualisations to 11-16 year old kids. Have trained 22 kids in Power BI

Agenda

1. What is Web Scraping ?
2. Web Scraping Use Cases
3. Web Scraping Simple
4. Get data by example
5. Get Data from HTML/CSS
6. Web Scraping using Power Query

What is Web Scraping

- Web Scraping is a process used to extract data from websites

What is Web Scrapping used for ?

- Build a price comparison website
- Build a product catalog
- Competitor Analysis
- Financial data analysis
- Lead Generation
- Brand Monitoring

What is Web Scrapping used for ?

- Gathering data for Machine Learning
- Displaying job boards with job listings
- MAP compliance monitoring
- Social media analysis
- News monitoring
- Search Engine Optimisation

Web Scraping Simple

- Data Source:
- [https://en.wikipedia.org/wiki/List_of_countries_by_population_\(United_Nations\)](https://en.wikipedia.org/wiki/List_of_countries_by_population_(United_Nations))
- Use the Web Connector after clicking the Get Data button



Get data by example

- Data Source:
- <https://www.metservice.com/maps-radar/local-observations/local-3-hourly-observations>
- References:
- <https://docs.microsoft.com/en-us/power-bi/desktop-connect-to-web-by-example>

Get Data from HTML/CSS

- Data Source:
- <https://radacad.com/blog>
- References:
- <https://www.amazon.com/Power-BI-MVP-Book-techniques-ebook/dp/B07WHC4PFV> -- 2nd
Chapter from this book

Webscraping using Power Query

- Data Source:
<https://www.quanthockey.com/khl/seasons/2017-18-khl-players-stats.html>
- Using Fiddler
- [http://www.quanthockey.com/scripts/AjaxPaginate.php?cat=Season&pos=Players&SS=2017-18&af=0&nat=2017-18&st=reg&sort=P&so=DESC&page="&Text.From\(PageNumber\)&"&league=KHL&lang=en&rnd=276273473&dt=1](http://www.quanthockey.com/scripts/AjaxPaginate.php?cat=Season&pos=Players&SS=2017-18&af=0&nat=2017-18&st=reg&sort=P&so=DESC&page=)

Webscraping using Power Query

- New Parameter for PageNumber
- Parameter
- (PageNumber as number) => let
- Source =
Web.Page(Web.Contents("http://www.quanthockey.com/scripts/AjaxPaginate.php?cat=Season&pos=Players&SS=2017-18&af=0&nat=2017-18&st=reg&sort=P&so=DESC&page=" & Text.From(PageNumber) & "&league=KHL&lang=en&rnd=276273473&dt=1")),

Webscrapping using Power Query

- All Pages code
- = List.Generate()=>
- [Result= try PageData(1) otherwise null, Page = 1],
- each [Result] <> null,
- each [Result= try PageData(Page) otherwise null, Page = [Page] + 1],
- each [Result])

Webscrapping using Power Query

- References:
- <https://www.sumproduct.com/thought/referencing-multi-page-data-from-a-web-site-with-a-single-url>

Recap of what we learnt

1. What is Web Scraping ?
2. Web Scraping Use Cases
3. Web Scraping Simple
4. Get data by example
5. Get Data from HTML/CSS
6. Web Scraping using Power Query