



IBM CAPSTONE PROJECT – THE BATTLE OF NEIGHBORHOODS

Cluster Analysis of London Real Estate Market

A low-angle, black and white photograph of several tall skyscrapers reaching towards a cloudy sky. The perspective is from the ground looking up, creating a sense of height and scale. The buildings are modern with many windows. A large, semi-transparent white rectangle is overlaid on the left side of the image, containing the text. A green rectangular area is also visible, partially overlapping the white rectangle and the buildings.

BUSINESS PROBLEM SECTION

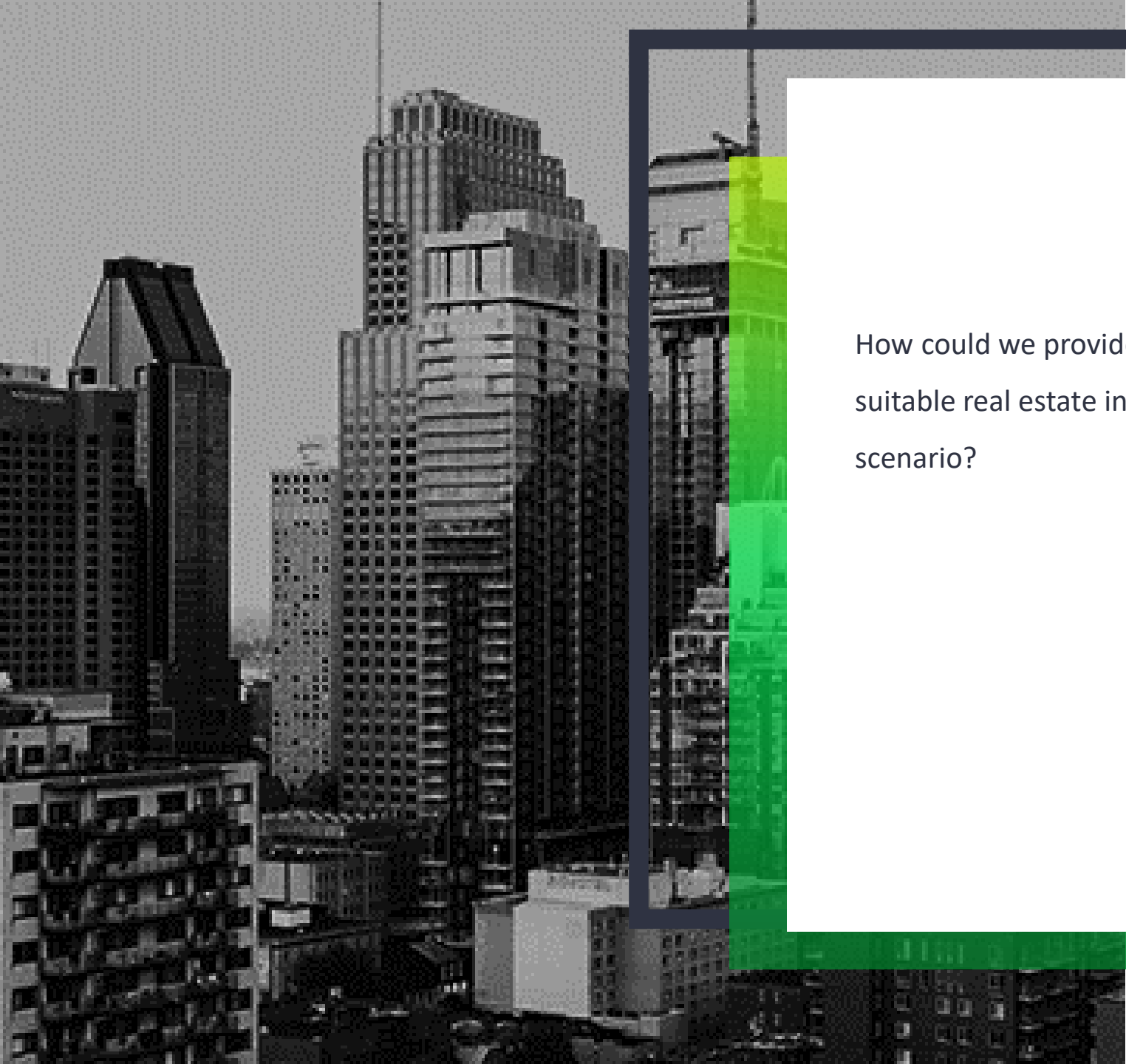
London Housing Market is in a rut:

1. Brexit
2. Hidden price falls
3. Record-low sales
4. Homebuilder exodus
5. Tax hikes addressing overseas buyers of homes in England and Wales.



The background is a grayscale aerial photograph of a dense urban area with numerous skyscrapers. Overlaid on this is a graphic consisting of a green rectangle with a white square inside it. A thick black L-shaped line frames the white square, starting from the top-left corner of the green rectangle and ending at the bottom-right corner of the white square.

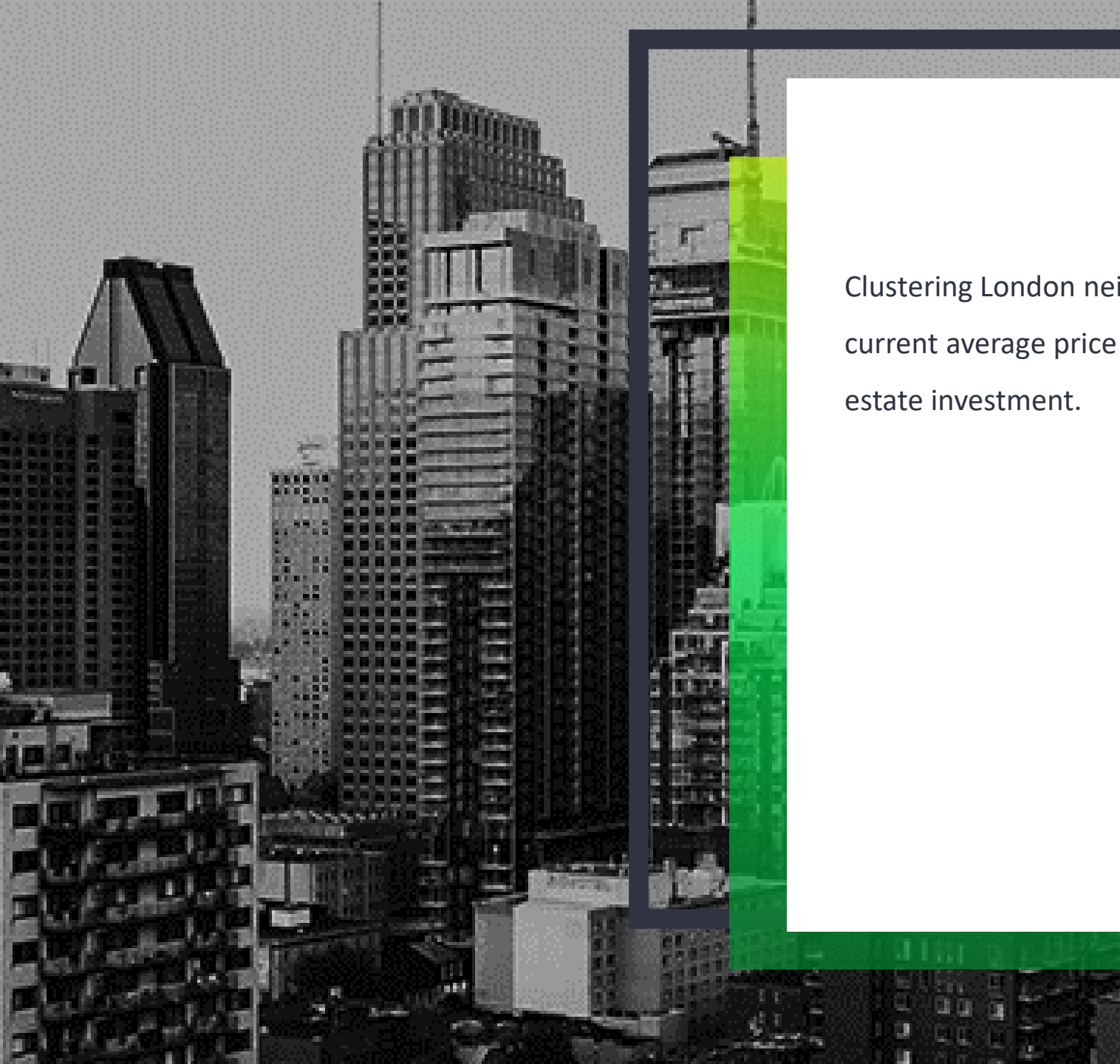
BUSINESS PROBLEM



How could we provide support to homebuyers clientele in to purchase a suitable real estate in London in this uncertain economic and financial scenario?

The background is a grayscale aerial photograph of a dense urban area with numerous skyscrapers. Overlaid on this is a graphic consisting of a green rectangle with a white square inside it. The white square has a thick black border. The word "SOLUTION" is centered within the white square.

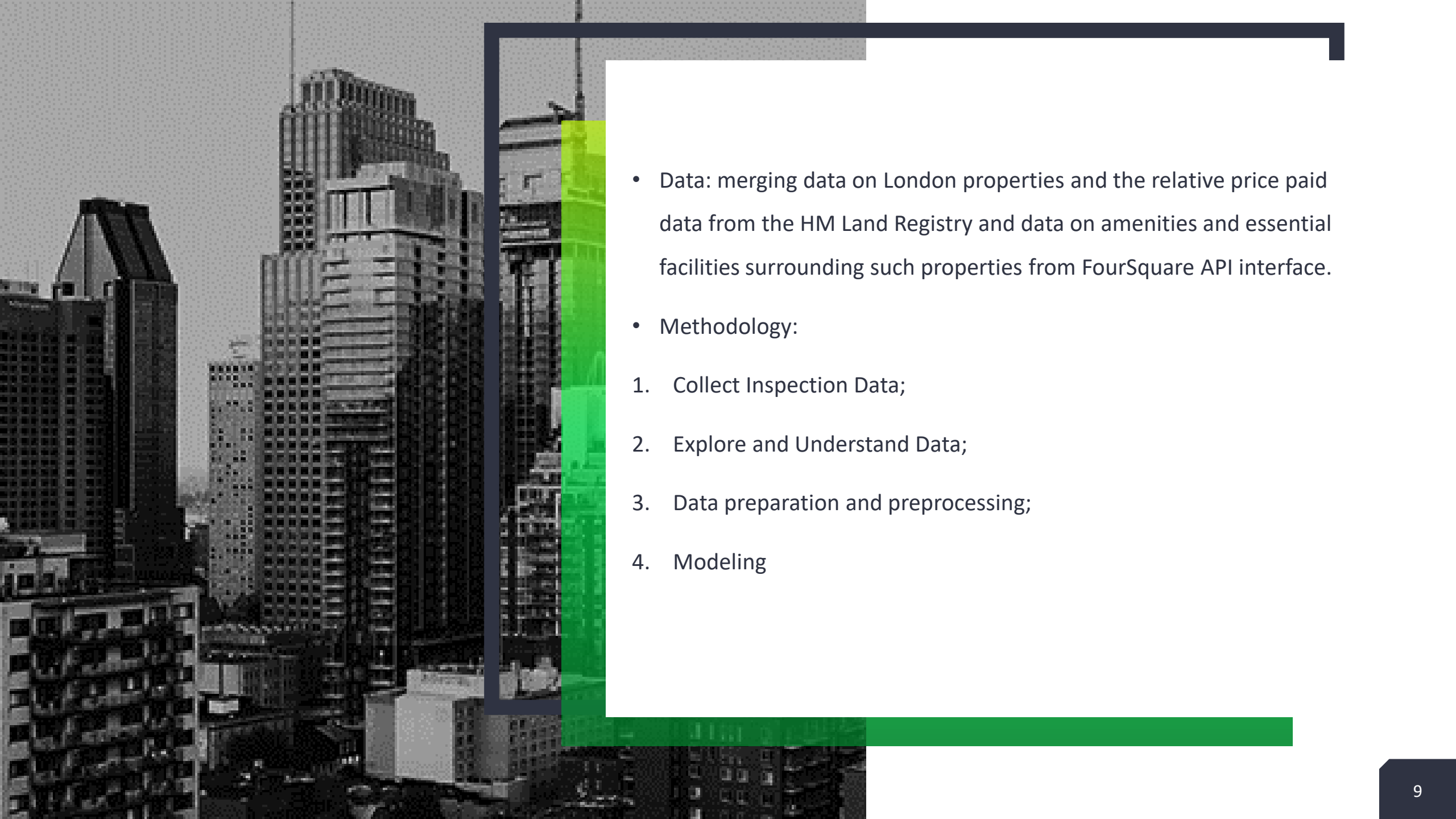
SOLUTION

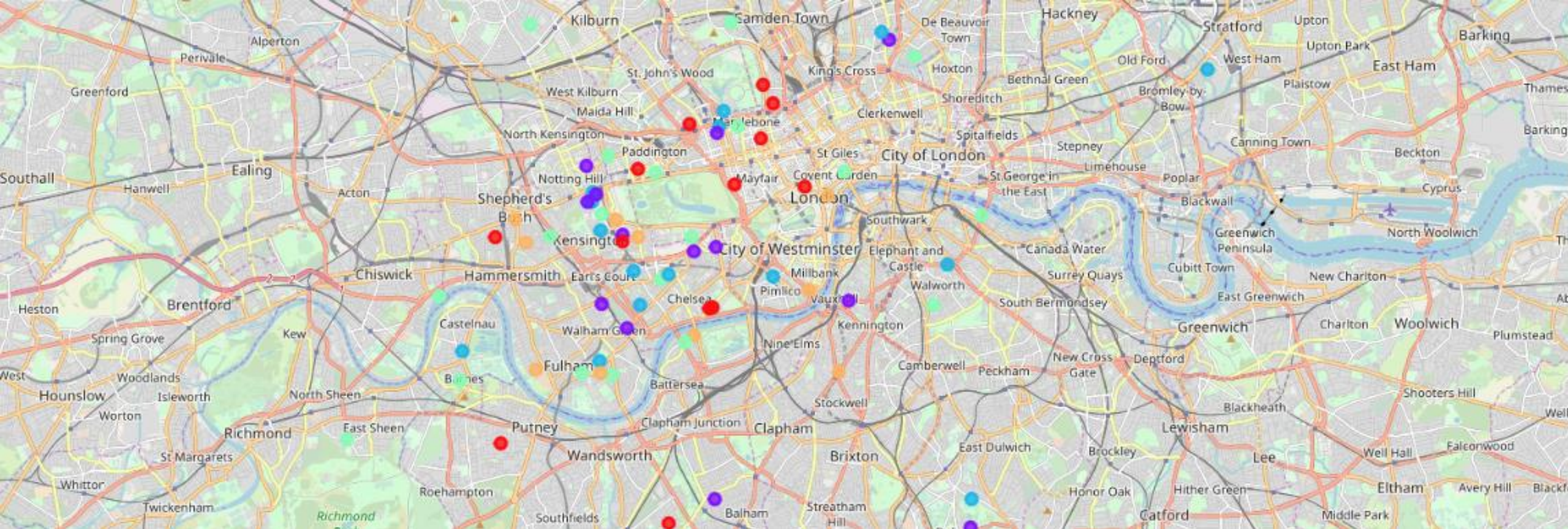


Clustering London neighborhoods in order to recommend venues and the current average price of real estate where homebuyers can make a real estate investment.

The background is a grayscale aerial photograph of a dense urban area with numerous skyscrapers. Overlaid on this is a graphic consisting of a green rectangle on the left and a white rectangle on the right, separated by a dark blue L-shaped line. The text 'DATA AND METHODOLOGY' is centered in the white rectangle.

DATA AND METHODOLOGY

- 
- Data: merging data on London properties and the relative price paid data from the HM Land Registry and data on amenities and essential facilities surrounding such properties from FourSquare API interface.
 - Methodology:
 1. Collect Inspection Data;
 2. Explore and Understand Data;
 3. Data preparation and preprocessing;
 4. Modeling



K-Means clustering

The background is a grayscale aerial photograph of a city with numerous skyscrapers. Overlaid on this is a graphic consisting of a green rectangle on the left and a white rectangle on the right, separated by a thick black L-shaped line. The word 'OUTCOME' is centered in the white rectangle.

OUTCOME



- Examination of real estates according to neighborhoods/London areas
 1. West London (Notting Hill, Kensington, Chelsea, Marylebone) and North-West London (Hampsted) might be considered highly profitable venues to purchase a real estate;
 2. South-West London (Wandsworth, Balham) and North-West London (Islington) are arising as next future elite venues with a wide range of amenities and facilities.
- Examination of real estates by clusters
 1. Clusters 0, 2 and 4 may target home buyers prone to live in 'green' areas with parks, waterfronts;
 2. Clusters 1 and 3 may target individuals who love pubs, theatres and soccer.



THANK YOU