

Exploratory Data Analysis and Visualisation of UK oil production during the period 1980 to 2010

Submitted By:

Rahul Chahel
52100749

Submitted to:

Dr. Shaktiman Singh
Course Coordinator-GG5065

Executive Summary

Geo-spatial analysis of UK offshore oil production during 1980-95 and 1995-2010 provided significant evidence of various patterns.

Presence and Distribution of wells by different factors like Geological basin, Cartographic quadrant and water depth was evidenced from the analysis.

The Discovery of new wells and change in production of oil from each well can be analysed from the available dataset.

The increment and decrement in the oil production for each well could be related to the Geographical Basin.

Statically, Scotland contributes nearly equal or more than rest of United Kingdom in terms of oil production.

Table of Content

Introduction.....	1
Distribution of Wells by Geological Major Basin.....	1
Distribution of wells by Cartographic Quadrant.....	3
Distribution of wells by Water Depth.....	5
Oil Production map.....	6
Period:1980-1995.....	6
Period: 1995-2010.....	7
Visual Pattern of Oil Production.....	8
Statistical Analysis of Oil Production at two periods.....	9
Pattern Change Analysis.....	12
Change in production analysis by each well.....	13
Analysis of Oil Production Change with other layers.....	15
Additional Analysis.....	17
Methodology.....	20
Conclusion.....	23

I. Introduction

Geo-spatial, Statistical analysis and Visualisation is exponentially growing methodology to analysis different datasets. There are several advantages to Geo-Spatial analysis as the requirement of refence frame for different data relation is fulfilled by Geological coordinates or other geological attributes.

Implementation of GIS in Sea related services like goods transfer, sea routes and sea resource management was there for long time. GIS implementation in oil industry along with other technologies like Machine learning and Artificial Intelligence is increasing rapidly. Similar exploratory data analysis and visualisation is presented in this report for UK oil production for time period 1980 to 2010.

II. Distribution of wells by Geological Major Basins

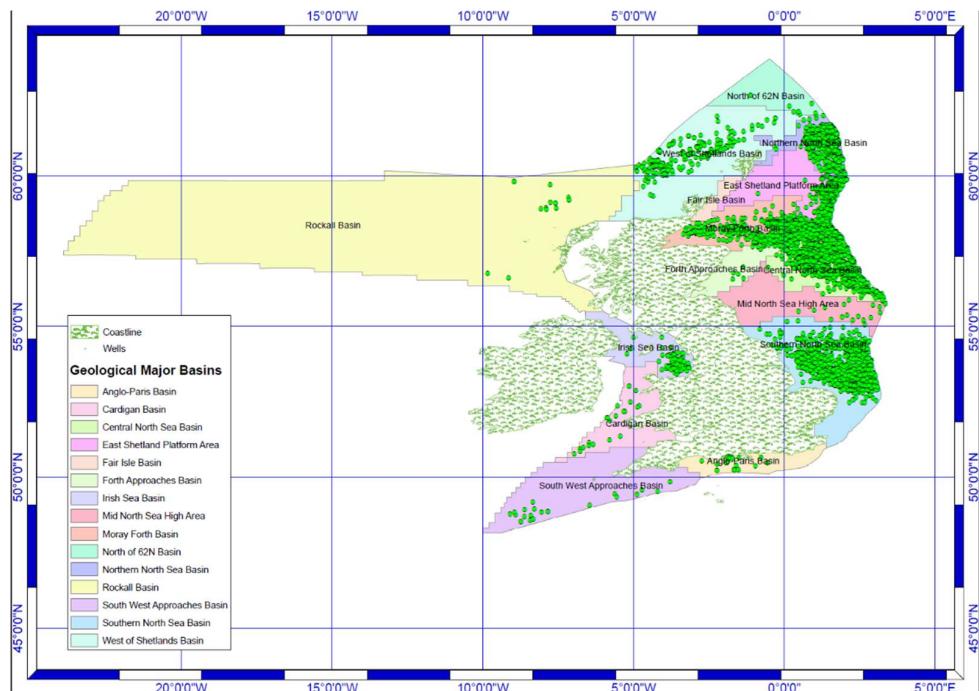


Figure 1 Distribution of Wells by Geological Basin (Map)

Map analysis of Figure 1 clearly shown pattern of heavy cluster of wells in certain Basins and lack of cluster in certain Basins. While Map analysis was perfect technique to get rough idea about the distribution of Wells by Basins, But for better quantitative analysis the Graphical visualisation embedded with quantitative terms to be adopted as shown in Figure 2.

The quantitative analysis evidenced that majority of wells are present in four basins i.e Northern North Sea Basin, Central North Sea Basin and Moray Firth Basin. Rockal Basin is the Basin with maximum area but only 15 wells are present there which is lowest Well per Km². Northern North Sea Basin is embedded with maximum number of wells and highest Wells per Km².

Table 1. Quantitative data of wells by Basin

Basin	Number of Wells	Area(Km ²)	Well per Km ²
Northern North Sea Basin	3694	34436	.1072
Central North Sea Basin	2349	46493.8	.050522
Southern North Sea Basin	2139	103168	.020733
Moray Forth Basin	2015	57251.5	.03519
West of Shetland Basin	454	113351	.004
East Shetland Platform	304	44805.6	.00678
Iris Sea Basin	268	39737	.006744
Anglo-Paris Basin	28	29638.7	.000944
Forth Approaches Basin	23	25047.5	.0009182
Cardigan Basin	22	50649.6	.0004343
Southwest Approaches Basin	22	92836.2	.000236977
Mid North Sea High Area	17	55861	.00030432
Rockall Basin	15	499648	.00003002
North of 62N Basin	9	32170.3	.000279
Fair Isle Basin	3	8089.91	.0003708

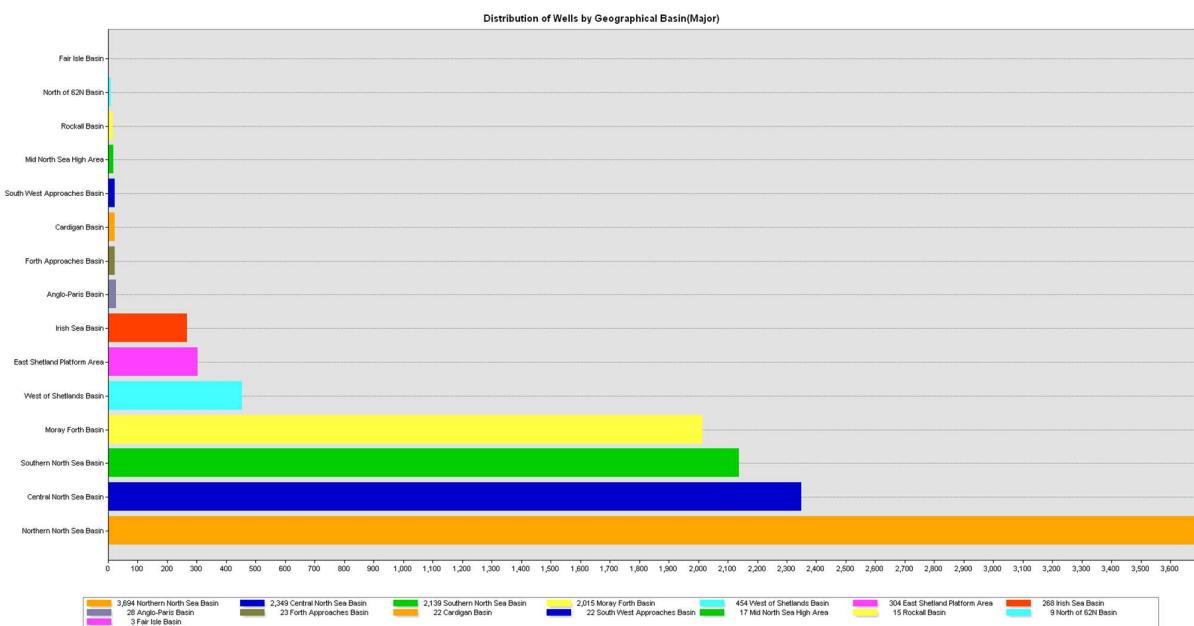


Figure 2 Distribution of Wells by Geographical Basin

III. Distribution of wells by cartographic quadrant

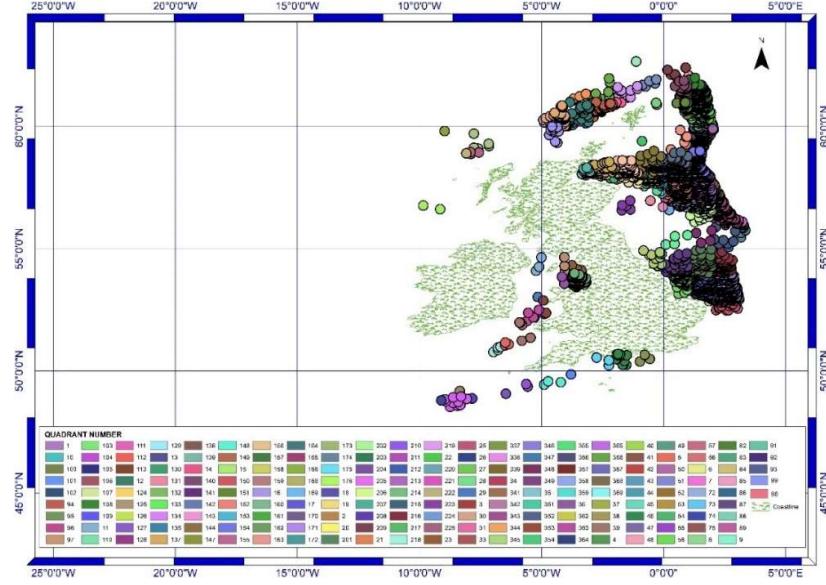


Figure 3 Distribution of Wells by Cartographic Quadrant (Map)

Map Analysis of wells by Cartographic Quadrant doesn't describe any pattern in the Zoomed out view, there is a possibility that it can provide any significant pattern in Zoomed in view as the area of quadrant is small for overall analysis. The Quantitative analysis of each quadrant some what shown pattern of cluster relation among neighbouring quadrants, but no assumption could be made based on quadrants individually related to frequency of wells.

Table 2. Distribution of Wells (Frequency) by Quadrant

Quad	FREQUENCY	Quad	FREQUENCY	Quad	FREQUENCY	Quad	FREQUENCY
1	2	27	2	73	11	132	2
2	142	28	31	74	2	153	1
3	663	29	186	83	1	154	3
4	3	30	513	85	1	163	2
7	1	31	44	86	2	164	7
8	8	36	4	87	3	202	10
9	883	37	4	88	1	204	229
10	31	38	9	93	4	205	57
11	68	39	11	97	4	206	91
12	47	41	17	98	19	207	7
13	327	42	125	99	5	208	15
14	314	43	147	102	3	209	6
15	728	44	206	103	5	210	158
16	980	47	189	106	5	211	1354
18	3	48	507	107	3	213	19
19	13	49	825	108	2	214	18
20	192	50	14	109	2	217	2
21	801	52	18	110	225	219	6
22	730	53	82	111	4	220	2
23	151	54	8	112	3		
26	5	72	1	113	34		

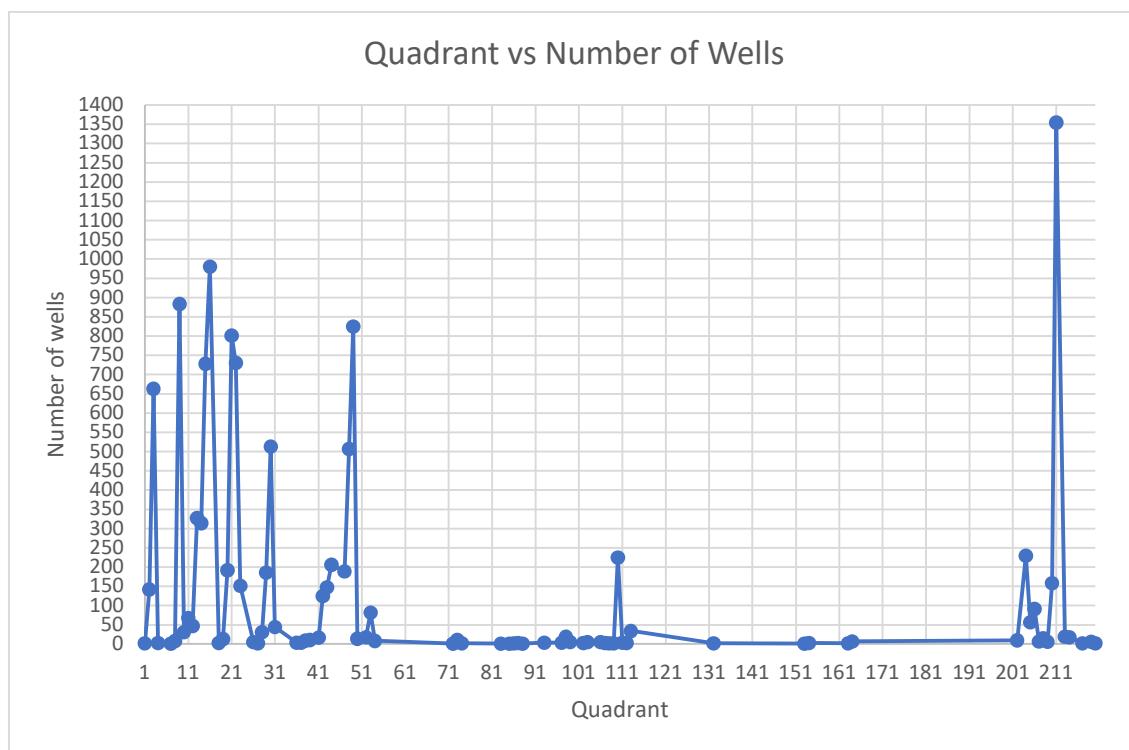


Figure 4 Graphical Representation of Quadrants vs Number of wells

IV. Distribution of wells by Water Depth

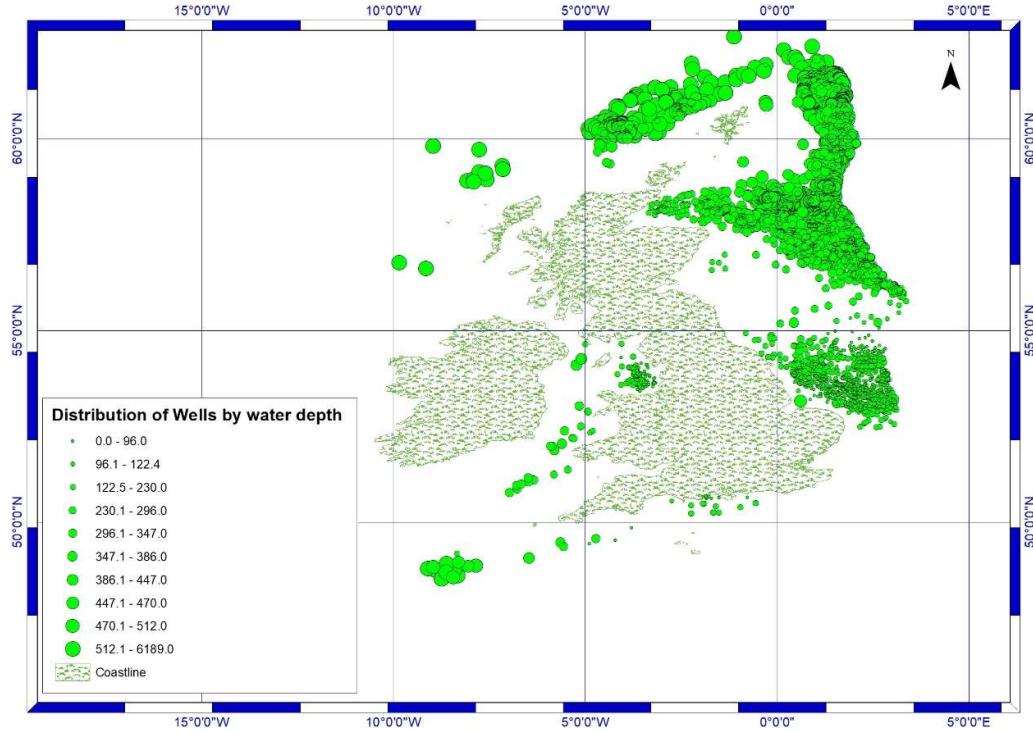


Figure 5 Distribution of Wells by water depth

Map analysis shown a significant presence of wells in cluster at less depth compare to more depth. Oil production relation with the water depth can be determined by the data but by this Map analysis it could be assumed that there is a inverse relation of number of wells and water depth.

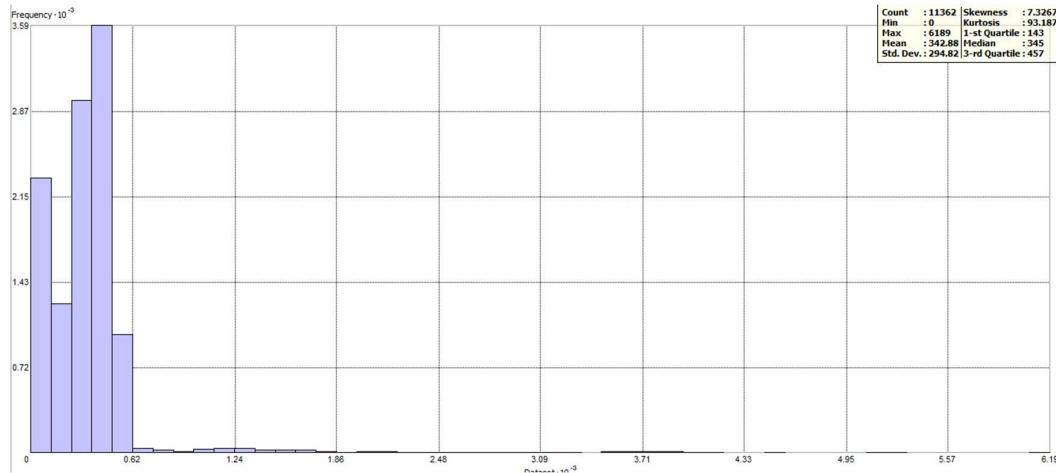


Figure 6 Histogram representing Distribution of Wells by water depth

The Statistical analysis by plotting Histogram also signify that more number of wells are present at less depth. The bell curve is highly shifted towards the low values which signify that median is too less than the mean. 1-st Quartile is 143 which signify that 1/4th of 11362 wells are present at less water depth than 143 and 3rd Quartile is 457 which signigy that 3/4th of 11362 wells are present at less water depth than 457.

V. Oil Production map

Period: 1980-1995

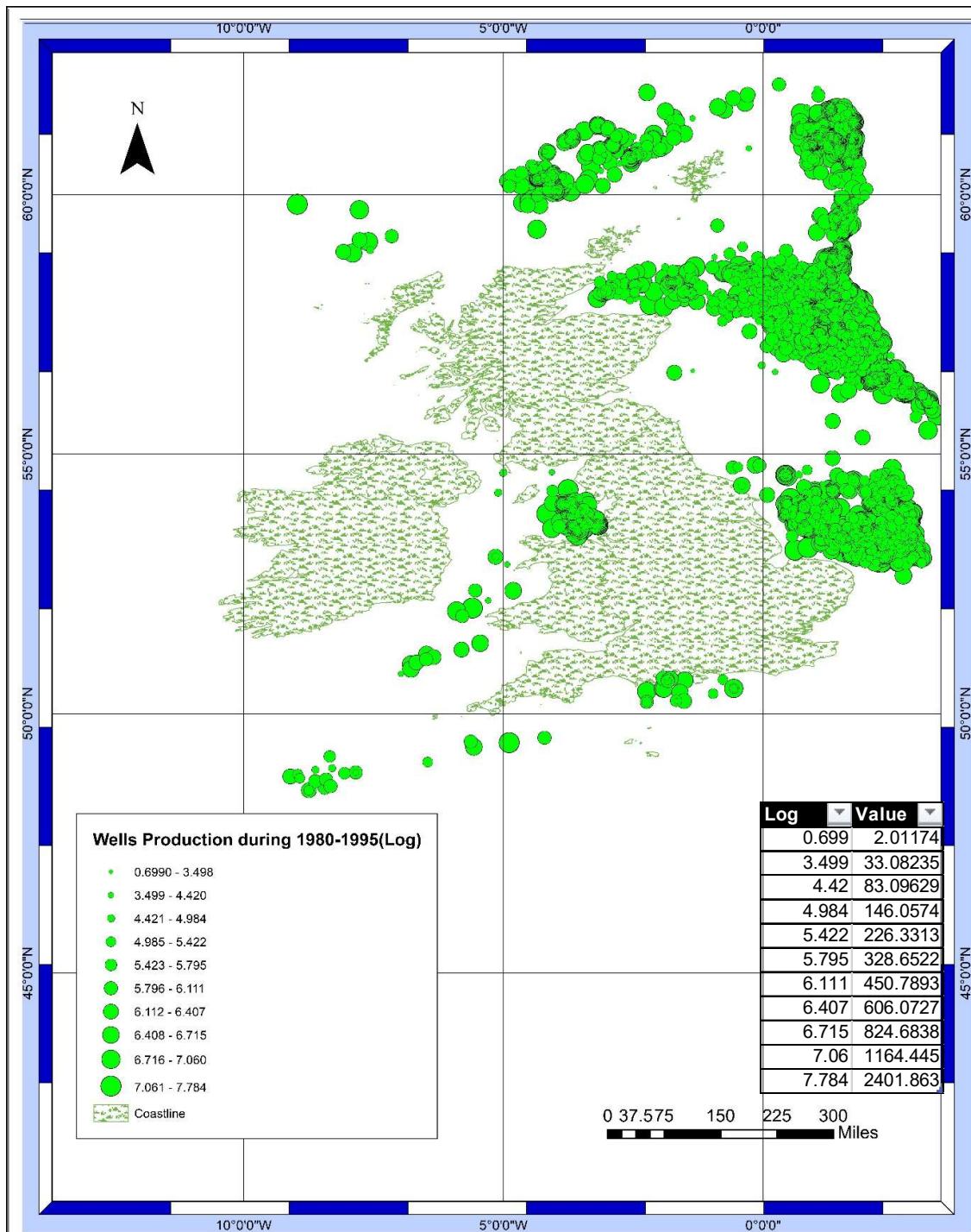


Figure 7 Oil production during 1980-1995 (by well)

Period: 1995-2010

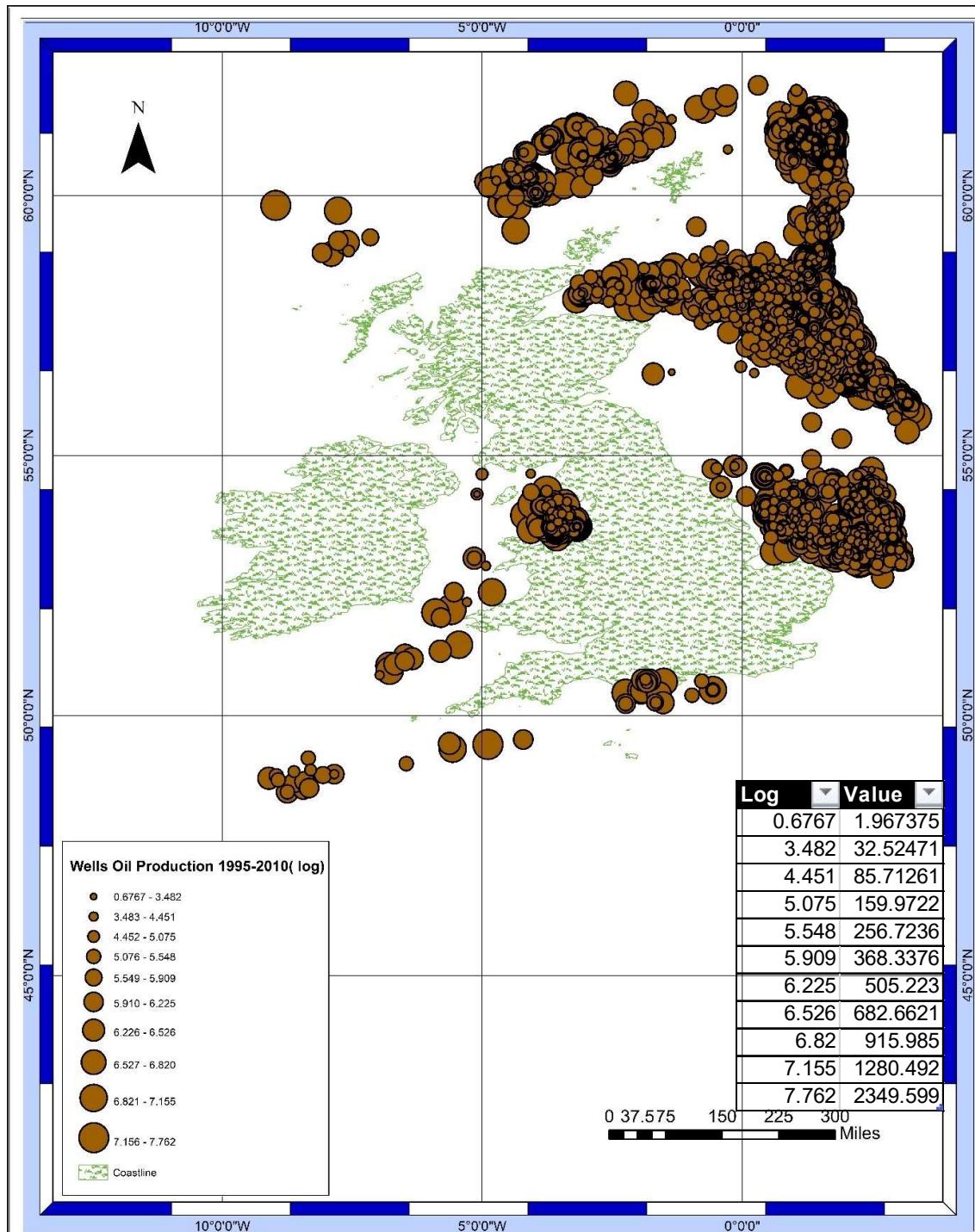


Figure 8 Oil production during 1995-2010 (by well)

VI. Visual Pattern of Oil Production

By Visual Map analysis of oil production during period 1980-95 and 1995-2010, the increment in number of wells could be assumed as the change in cluster of wells in positive by appearance. The oil production can also be considered as increased because the legend scale represent higher amount of oil production and number of big sized well symbology is dominating in 1995-2010 compare to 1980-95.

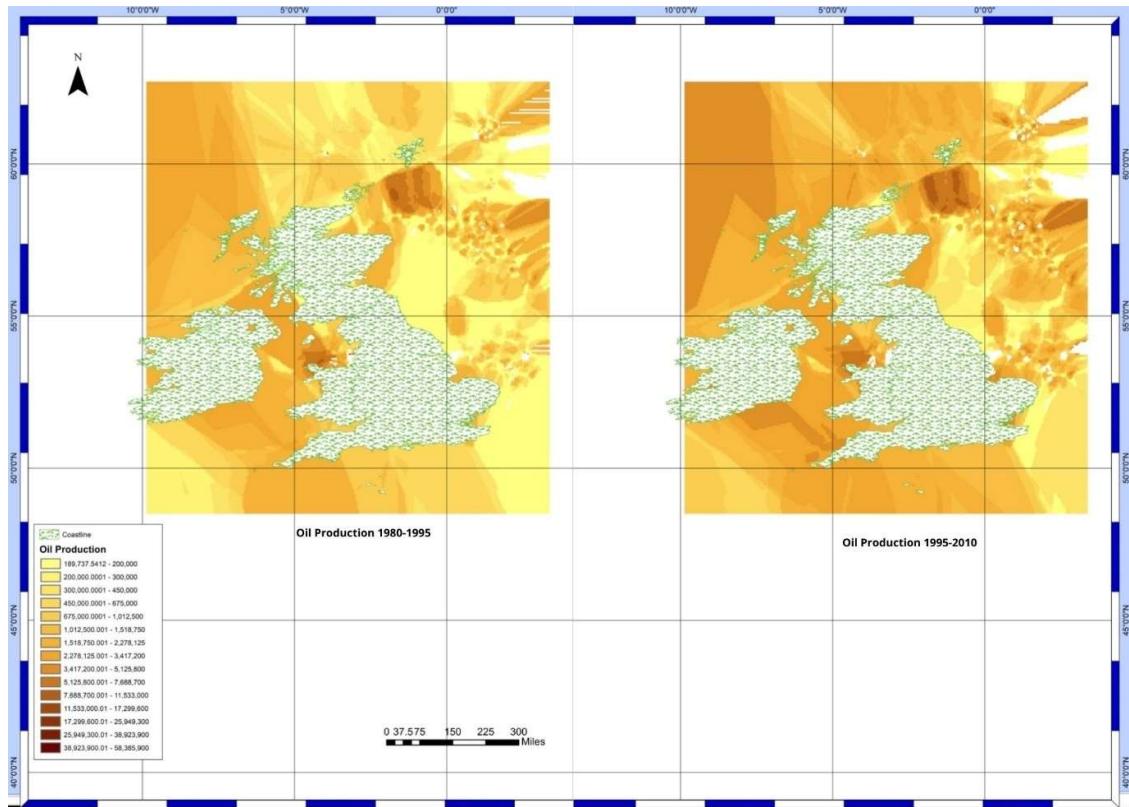


Figure 9 Kriging representation of Oil production during 1980-95 and 1995-2010

In case of Kriging comparison for both time period's oil production on same scale, it can be observed and concluded that oil production increased significantly in certain regions.

VII. Statistical Analysis of oil production at two periods

A. Before Pre-processing (Cleaning data)

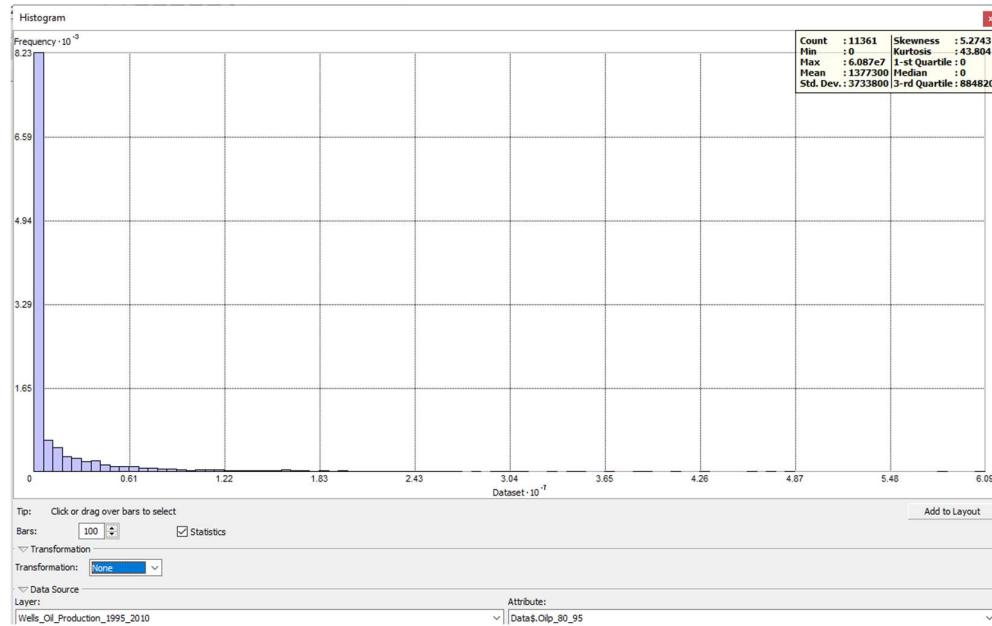


Figure 10 Statistical Analysis of Oil production for time period 1980-95 (before Pre-processing)

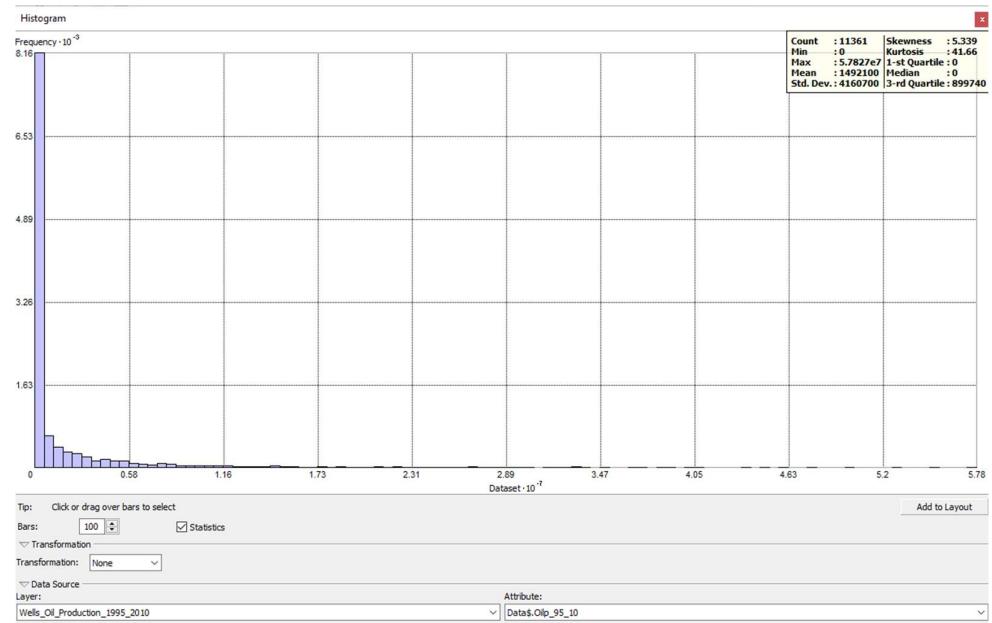


Figure 11 Statistical Analysis of Oil production for time period 1995-2010 (before Pre-processing)

Before Pre-processing the 1-st Quartile and Median of oil production for both time period was 0, that signify that there large quantity of null values or data with 0 present in the attribute table. It could be due to error in data source or due to operations during analysis. To resolve the issue new attribute table was exported with all data with oil production greater than 0 (Period 1 OR Period 2).

B. After Pre-Processing

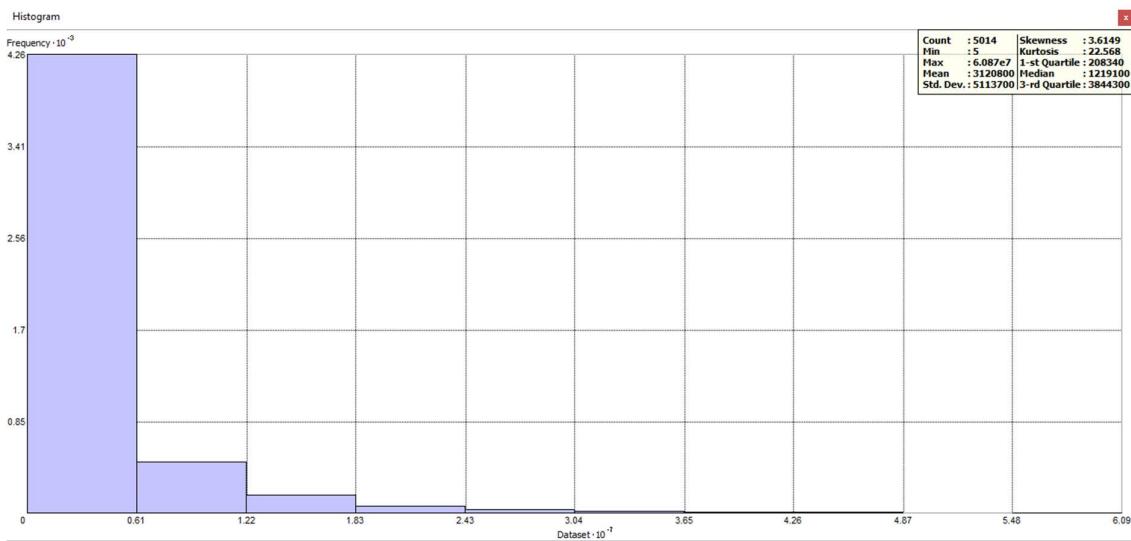


Figure 12 Statistical Analysis (with Histogram) of Oil production for time period 1980-95

For production period 1980-95, minimum oil produced by a well was 5 and maximum oil produced by a well was 60870000. 25% of wells, i.e. 1st Quartile, produced less than 208340 (inclusive) while 25% of wells, i.e. 4th Quartile, produced more than 3844300(exclusive). Median was 1219100, which represent that 50% of wells produced oil more than that value and 50% of wells produced less than that value. Other percentages can be calculated approximately using empirical rule formula as standard deviation is known. Mean is too higher compare to median, it represent that the curve is asymmetric bell curve with too many observations on lower segment.

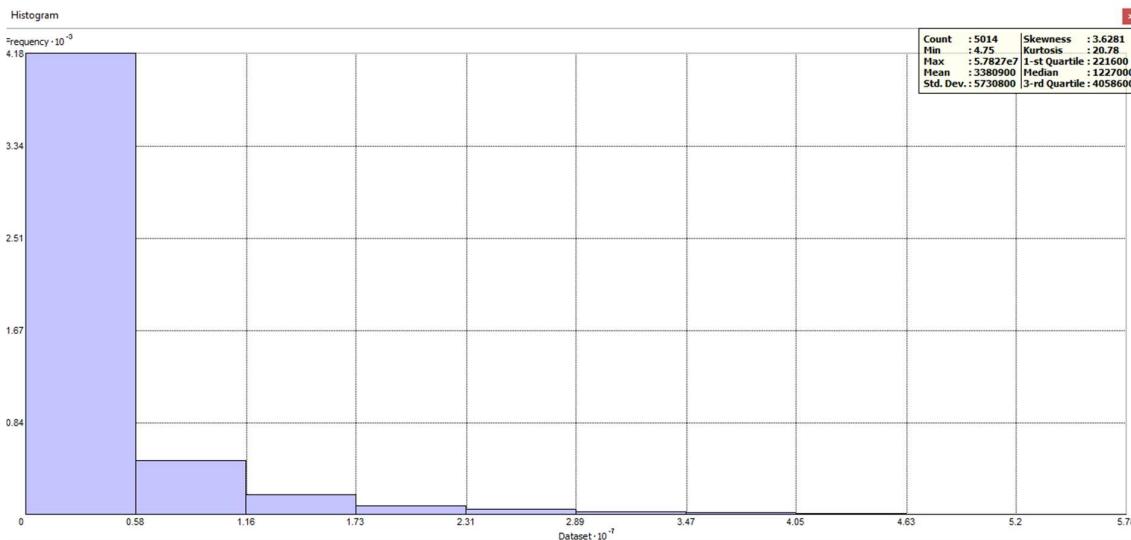


Figure 13 Statistical Analysis (with Histogram) of Oil production for time period 1995-2010

For production period 1980-95, minimum oil produced by a well was 4.75 and maximum oil produced by a well was 57827000. 25% of wells, i.e. 1st Quartile, produced less than 221600 (inclusive) while 25% of wells, i.e. 4th Quartile, produced more than 4058600(exclusive). Median was 1227000, which represent that 50% of wells produced oil more than that value and 50% of wells produced less than that value. Other percentages can be calculated approximately using empirical rule formula as standard deviation is known. Mean is too higher compare to median, it represent that the curve is asymmetric bell curve with too many observations on lower segment.

Table 3. Quantitative facts related to oil production during period 1980-95 and 1995-2010

Column1	Production period 1(1980-95)		Production Period 2 (1995-2010)	Change
Count		5014	5014	
Min		5	4.15	-0.85
Max		60870000	57827000	-3043000
Mean		3120800	3380900	260100
Std. Dev.		5113700	5730800	617100
Skewness		3.6149	3.6281	0.0132
Kurtosis		22.568	20.78	-1.788
1-st Quartile		208340	221600	13260
Median		1219100	1227000	7900
3-rd Quartile		3844300	4058600	214300

On subtraction of Production period 1 statistics with Production Period 2 Statistics, The significant rise of oil production can be observed. There was decrease in well will maximum production but change in all three Q1(1st Quartile), Q2(Median) and Q3(3rd Quartile) was positive. This positive change signify that the oil production in 1995-2010 was more and improved rate compare to oil production in 1980-1995.

VIII. Pattern Change Analysis

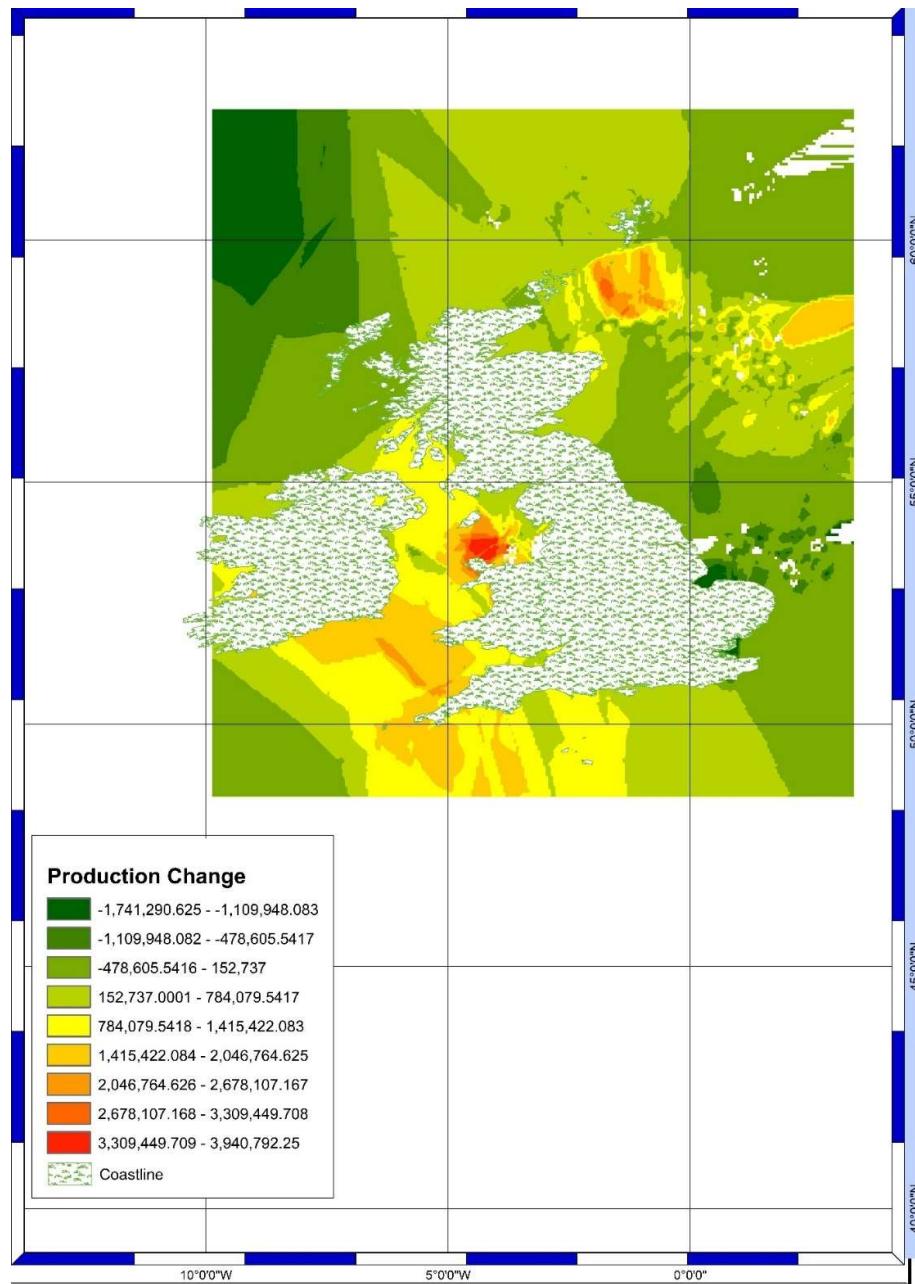


Figure 14 Production Change Kriging by well

Map Analysis of Oil production change Kriging give a mirage of negative change in oil production as comparatively more area on the map is covered with colour representing negative change. But the amplitude of positive change is more and legends representing positive change are in majority.

There are certain areas on map that display high positive change in oil production, i.e. Area of sea present between 5°W-0°W and 50°N-60°N.

IX. Change in production of each well

A. Wells with negative change in oil production

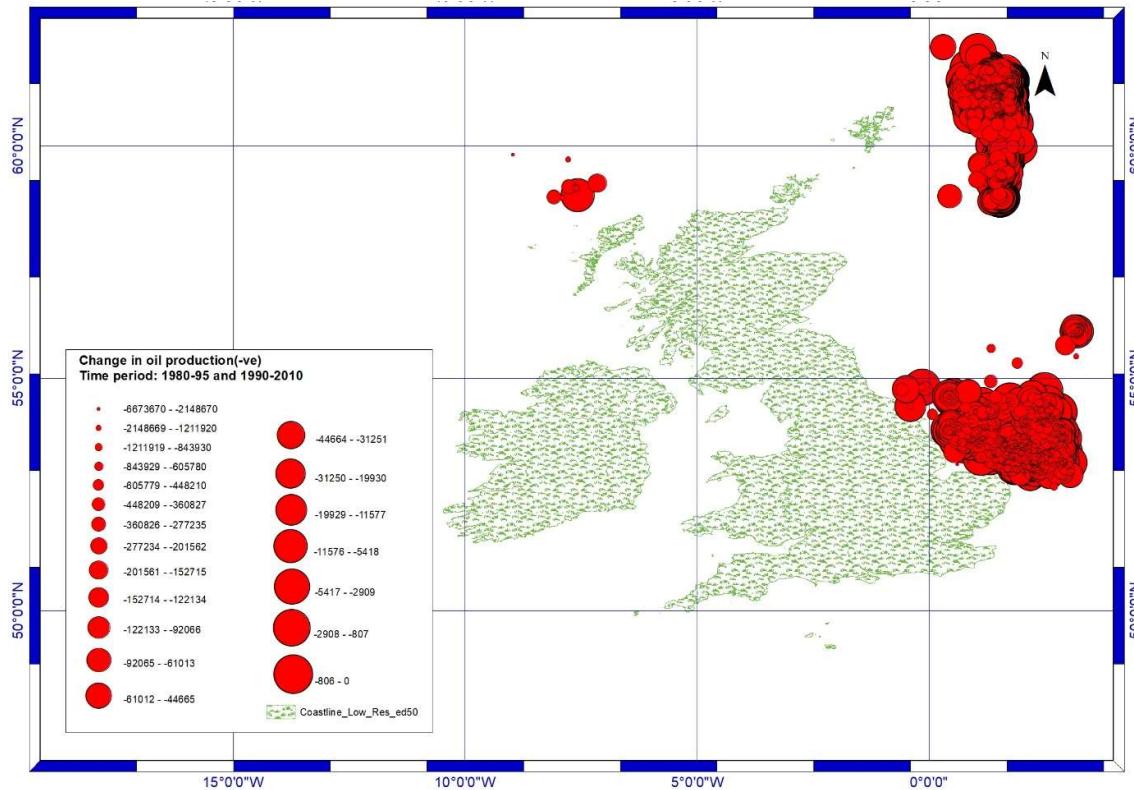


Figure 15 Change in Oil Production between Time Period 1980-95 and 1995-2010 by well (Negative)

B. Wells with Positive change in oil production

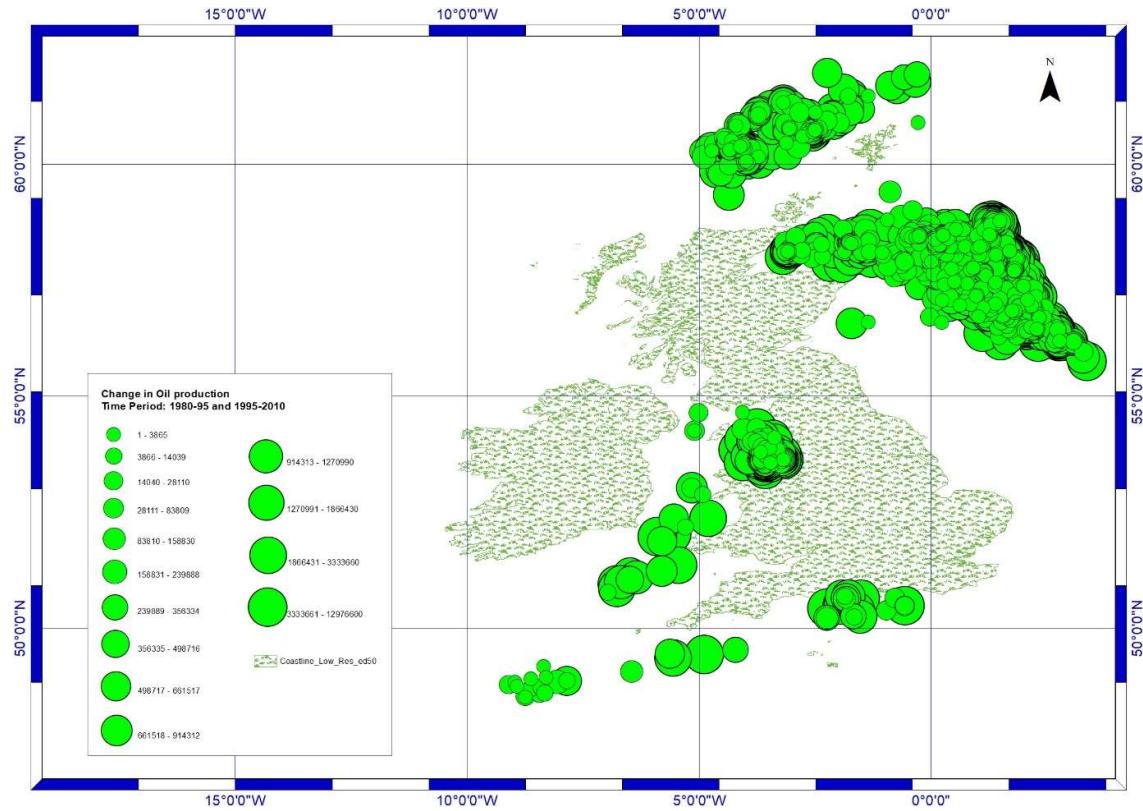


Figure 16 Change in Oil Production between Time Period 1980-95 and 1995-2010 by well (Positive)

X. Analysis of Oil Production change

By Map analysis of production change of each well with positive and negative as different layer, the cluster of similar category can be observed. To relate the similarity with geological perimeter, layer with major basin displayed significant relation to the change in oil production layer.

A. Positive Oil production change

The ‘Positive change oil production’ layer relation to basin layer signify that wells belonging to West of Shetlands Basin, Moray Forth Basin, Central North Sea Basin, Forth Approached Basin, Irish Sea Basin, Cardigan Basin, Anglo-Paris Basin and South West Approaches Basin had positive change in contribution to oil production.

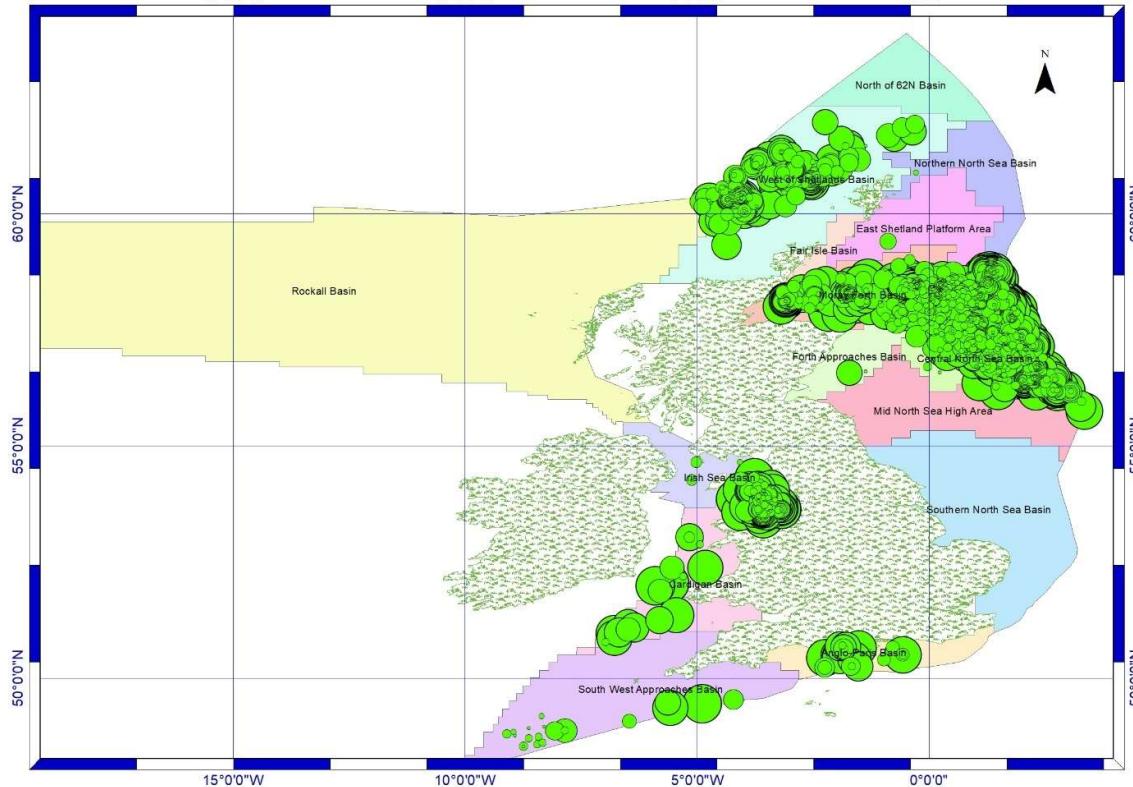


Figure 17 Positive Oil production change by well and Basin layer

B. Negative Oil production

The ‘Negative change oil production’ layer relation to basin layer signify that wells belonging to Rockall Basin, Northern North Sea Basin, Mid North Sea High Area, Southern North Sea Basin had negative change in contribution to oil production.

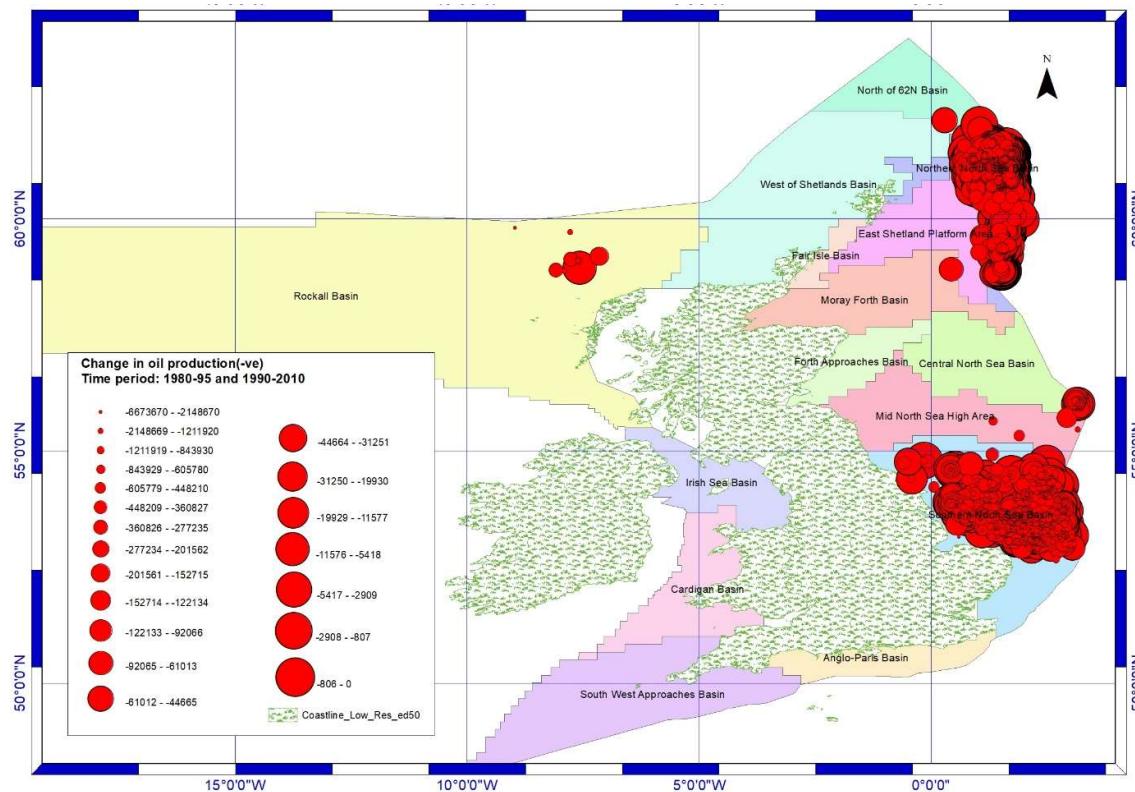


Figure 18 Negative Oil production change by well and Basin layer

XI. Additional Analysis

A. Blocks with potential good oil source

In the analysis of oil production change with respect to Basin, significant growth was present in certain basins. In layer ‘28rd offer’ there are number of blocks that are available to explore. Union of ‘28rd offer’ layer with basin layer’s positive production basins provide the data related to potential good oil source.

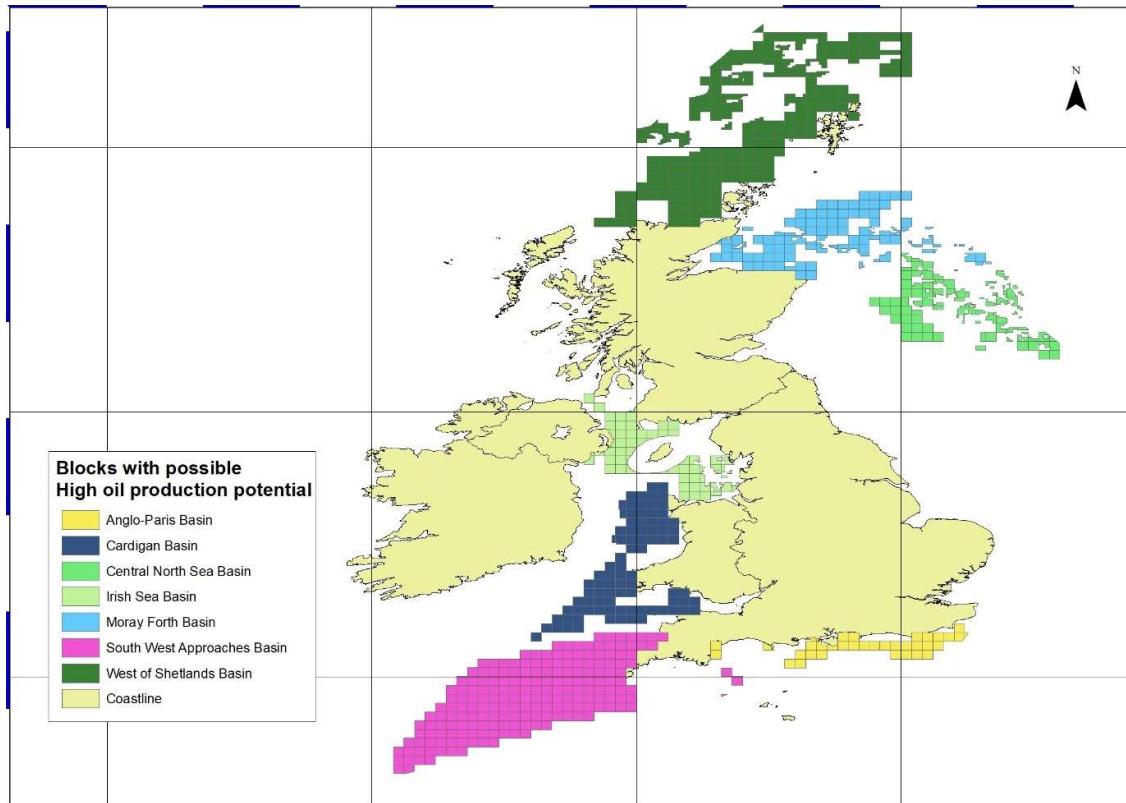


Figure 19 Blocks with Potential Good Oil presence

B. Analysis of Market share of different companies in different production field

By joining data of current operator of different well with type of production a well is suitable for, statistical data related to market shares of different operators in different production business could be generated with it.

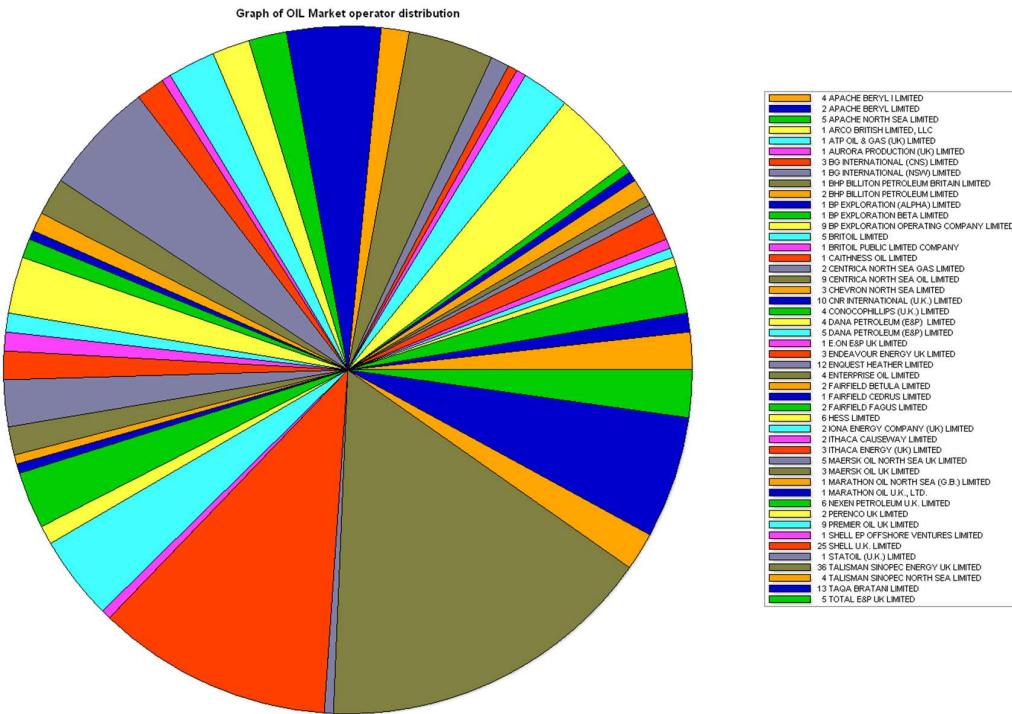


Figure 20 Operators with number of wells under control and used for Oil production

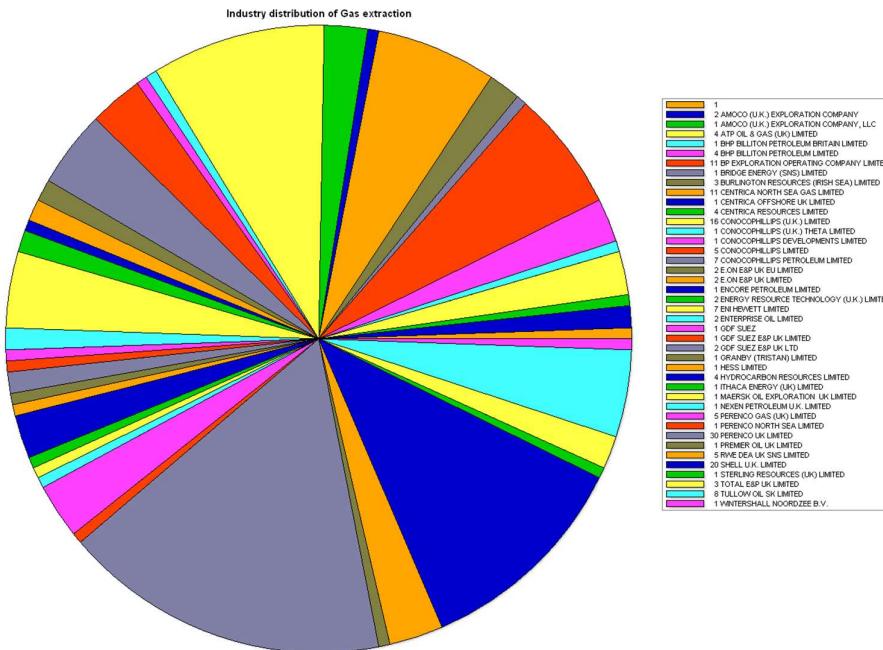


Figure 21 Operators with number of wells under control and used for Gas production

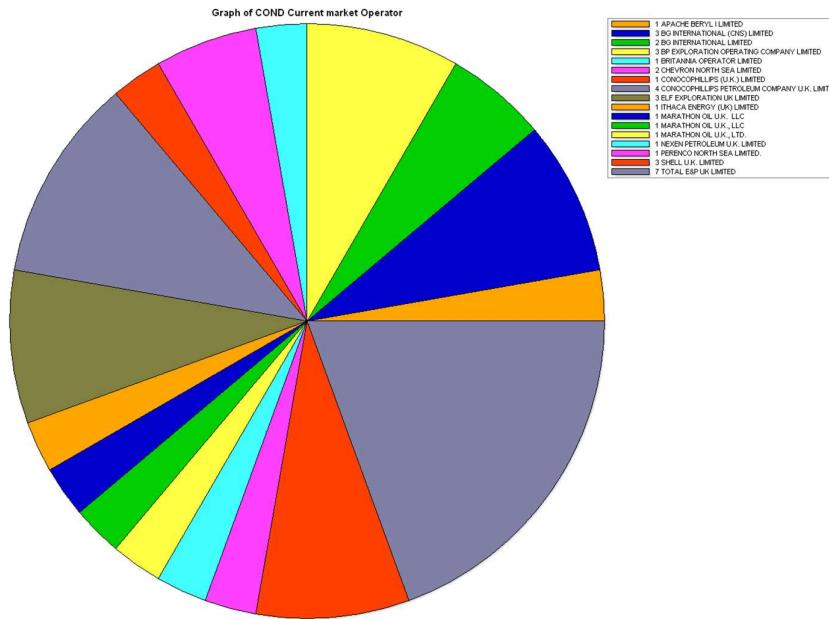


Figure 22 Operators with number of wells under control and used for Condensate production

Further relation of the well registration number of well with current operator data, it will generate the total production of different fuels by different operators.

C. Country/ County wise Oil production Analysis

Different part of UK contribute different amount of oil production. It can be analysed by Map reading of Basin Vs wells that Scotland produce nearly 50% or more oil. If layer with International boundaries or sea territorial boundaries can be included then better analysis of County wise oil production could be possible.

XII. Methodology

To map and analyse distribution of wells by Geological Basin

1. Insert Layer with Well data and Data with Geological Basin(Major or sub).
2. In ‘Table of Contents dialogue box’, Right click on layer with well data and Go to ‘Joins and Relates’ and Select Join.
3. Join Data Dialogue box will appear, Select ‘Join data from another layer based on spatial location’ then Choose layer with Basin data, i.e. ‘Geological_Basins_Major’ or ‘Geological_Basins_Sub’.
4. Select the polygon and point relation, i.e. if the point falls inside the polygon or is closest to it. In this case, Select ‘point falls inside’.
5. Choose the location of result layer by providing path in ‘Specify output shapefile or feature class for this new layer:’ textfield.
6. Click on ‘OK’. A new layer with layer will be generated with joined data of well layer and basin layer.
7. Add Symbology to the layer for better visualisation by editing ‘Layer Properties’. To access ‘Layer Properties’ Right click on Generated layer form ‘Table of Content’, Select Properties.
8. Go to Symbology Tab, select Show: ‘Categories’ then Unique Values. Select Coulm with Basin data as Value field . Choose Color Ramp and Add values of all the areas by clicking on Add All Values.
9. Click Ok to Visualise map with Categorized well data according to Basins. For Better Visualization turn on the Basin layer.
10. To create a Graphical visualisation of Number of wells vs Basin location, Go to attribute table of the layer, By right clicking on Layer name then ‘Open Attribute Table’, Click on Table options Button then Select ‘Create Graph’.
11. Create Graph Wizard will appear. Select Graph Type, Data Source Layer, Value Field and Data for each axis accordingly. Select Next, Enter details related to Title, Legend and other visual parameters. Click on Finish to generate Graph.

To map and analyse distribution of wells by Cartographic Quadrant

1. Insert Layer with Well data and Data with Quadrants.
2. In ‘Table of Contents dialogue box’, Right click on layer with well data and Go to ‘Joins and Relates’ and Select Join.
3. Join Data Dialogue box will appear, Select ‘Join data from another layer based on spatial location’ then Choose layer with Quadrant data, i.e. ‘Quadrant_All’.
4. Select the polygon and point relation, i.e. if the point falls inside the polygon or is closest to it. In this case, Select ‘point falls inside’.
5. Choose the location of result layer by providing path in ‘Specify output shapefile or feature class for this new layer:’ textfield.
6. Click on ‘OK’. A new layer with layer will be generated with joined data of well layer and basin layer.
7. Add Symbology to the layer for better visualisation by editing ‘Layer Properties’. To access ‘Layer Properties’ Right click on Generated layer form ‘Table of Content’, Select Properties.
8. Go to Symbology Tab, select Show: ‘Categories’ then Unique Values. Select Column with Quadrant data as Value field . Choose ‘Color Ramp’ and Add values of all the areas by clicking on ‘Add All Values’.
9. Click Ok to Visualise map with Categorized well data by Quadrants. For Better Visualization turn on the Quadrant layer or Go to Layout View and add Legend with Active layers as ‘Legend Items’.

10. To create a Graphical visualisation of Number of wells vs Cartographic Quadrant, Go to attribute table of the layer, By right clicking on Layer name then ‘Open Attribute Table’, Click on Table options Button then Select ‘Create Graph’.
11. Create Graph Wizard will appear. Select Graph Type, Data Source Layer, Value Field and Data for each axis accordingly. Select Next, Enter details related to Title, Legend and other visual parameters. Click on Finish to generate Graph.

To map and analyse distribution of wells by Water Depth

1. Insert Layer with Well data and Csv file with water depth data, by using ‘Add data’ button on standard toolbar.
2. In ‘Table of Contents dialogue box’, Right click on layer with well data and Go to ‘Joins and Relates’ and Select Join.
3. Join Data Dialogue box will appear, Select ‘Join data from a table’.
4. Select the field in the layer that the join will be based on, i.e. ‘WELLREGNO’ in this case, and select the table to join the layer with, i.e. ‘Well_production.csv’. Select the field in the table to base the join on i.e. ‘WELLREGNO’ and choose weather if you want to keep all records or just the matching records. For error check prior to join click on ‘Validate Join’ or to join the table click on ‘OK’.
5. Join will be added to the parent shapefile and it’s attribute table.
6. Add Symbology to the layer for better visualisation by editing ‘Layer Properties’. To access ‘Layer Properties’ Right click on Generated layer form ‘Table of Content’, Select Properties.
7. Go to Symbology Tab, select Show: ‘Quantities’ then ‘Graduated symbols’. Select Column with Water depth data as Value field. Choose ‘Template’, Natural Breaks, Symbol size and other parameters.
8. Click Ok to Visualise map with Categorized well data by water depth. For Better Visualization Go to Layout View and add Legend with Active layers as ‘Legend Items’.
9. To create a Graphical visualisation, Kriging’, of Wells vs Water Depth, Go to Toolboxes in Catalog dialogue box followed by ‘System Toolboxes’ > ‘3D Analyst Tools.tbs’ > ‘Raster Interpolation’ > ‘Kriging’ OR search Kriging from ‘Search’ Dialogue box by pressing ‘Ctrl+F’.
10. Kriging Dialogue box will appear. Select ‘Input point features’ as Parent layer and ‘Z value field’ as Water depth column. Enter path for Output surface raster, Select Kriging method and Semivariogram model according to requirements. Click ‘OK’ to generate layer with Kriging result. Insert legend for better understandability of map and plotted data.

Derivation of maps of oil production from data. representing production during 1980-1995 and 1995-2010 or Any other duration

1. Insert Layer with Well data and Csv file with production data, i.e. ‘Well_production’, by using ‘Add data’ button on standard toolbar.
2. In ‘Table of Contents dialogue box’, Right click on layer with well data and Go to ‘Joins and Relates’ and Select Join.
3. Join Data Dialogue box will appear, Select ‘Join data from a table’.
4. Select the field in the layer that the join will be based on, i.e. ‘WELLREGNO’ in this case, and select the table to join the layer with, i.e. ‘Well_production.csv’. Select the field in the table to base the join on i.e. ‘WELLREGNO’ and choose weather if you want to keep all records or just the matching records. For error check prior to join click on ‘Validate Join’ or to join the table click on ‘OK’.
5. Join will be added to the parent shapefile and it’s attribute table.
6. Add Symbology to the layer for better visualisation by editing ‘Layer Properties’. To access ‘Layer Properties’ Right click on Generated layer form ‘Table of Content’, Select Properties.

7. Go to Symbology Tab, select Show: ‘Quantities’ then ‘Graduated symbols’. Select Column with oil production data of year, 1980-95 and 1995-2010 in separate analysis, as Value field. Choose ‘Template’, Natural Breaks, Symbol size and other parameters.
8. Click Ok to Visualise map with oil production during selected time period. For Better Visualization, Go to Layout View and add Legend with Active layers as ‘Legend Items’.
9. To create a Graphical visualisation of Wells vs Oil Production, Go to attribute table of the layer, By right clicking on Layer name then ‘Open Attribute Table’, Click on Table options Button then Select ‘Create Graph’.
10. Create Graph Wizard will appear. To visualise both time period in same graph click on ‘Add’, Select ‘Add Series’. For both Series, Select Graph Type, Data Source Layer, Value Field and Data for each axis accordingly. Select Next, Enter details related to Title, Legend and other visual parameters. Click on Finish to generate Graph.
11. To create a Contour visualisation, Kriging’, of Wells vs production, Go to Toolboxes in Catalog dialogue box followed by ‘System Toolboxes’ > ‘3D Analyst Tools.tbs’ > ‘Raster Interpolation’ > ‘Kriging’ OR search Kriging from ‘Search’ Dialogue box by pressing ‘Ctrl+F’.
12. Kriging Dialogue box will appear. Select ‘Input point features’ as Parent layer and ‘Z value field’ as oil production with time period of choice. Enter path for Output surface raster, Select Kriging method and Semivariogram model according to requirements. Click ‘OK’ to generate layer with Kriging result. Insert legend for better understandability of map and plotted data.

Statistical Analysis of Oil Production Pattern:

1. Right click on the Upper Bar and Select Geostatistical Analyst.
2. From Geostatistical Analyst floating toolbar, Go to Explore Data > Histogram.
3. Select Layer with oil production data in Layer selection drop box. Select data containing oil production during desired period in ‘Attribute’ drop down box.
4. The Histogram of desired data will be generated in the dialogue box along with the Statistical data on top with Min, Max, Std. Dev, Mean and other statistical features.
5. The Shape of Histogram will be the analysing factor in case of Statistical Analysis as different pattern of histogram represent different characteristics of the data.
6. Change in Statistical parameters of the data from different production year will reflect the change in global production.

Calculation of Change in production of each well and derivation of map

1. Insert Layer with Well data and Csv file with production data, i.e. ‘Well_production’, by using ‘Add data’ button on standard toolbar.
2. In ‘Table of Contents dialogue box’, Right click on layer with well data and Go to ‘Joins and Relates’ and Select Join.
3. Join Data Dialogue box will appear, Select ‘Join data from a table’.
4. Select the field in the layer that the join will be based on, i.e. ‘WELLREGNO’ in this case, and select the table to join the layer with, i.e. ‘Well_production.csv’. Select the field in the table to base the join on i.e. ‘WELLREGNO’ and choose whether if you want to keep all records or just the matching records. For error check prior to join click on ‘Validate Join’ or to join the table click on ‘OK’.
5. Join will be added to the parent shapefile and it’s attribute table.
6. Open the Attribute table of parent shapefile, Go to ‘Table Options’ and select ‘Add field’. Enter required details in Add Field dialogue box and select data type to ‘float’. Click ‘OK’ to add the field.
7. Right click on the column name of newly added field, Select ‘Field Calculator’ and enter the calculating operation in the blank text field with selecting fields from the ‘Fields’ text field.

Example: [oil_1980_95]-[oil_1995_10]. Click on ‘OK’ to automatically add the values to the field.

8. Add Symbology to the layer for better visualisation by editing ‘Layer Properties’. To access ‘Layer Properties’ Right click on Generated layer form ‘Table of Content’, Select Properties.
9. Go to Symbology Tab, select Show: ‘Quantities’ then ‘Graduated symbols’. Select Column with oil production change, newly added field, as Value field. Choose ‘Template’, Natural Breaks, Symbol size and other parameters.
10. To create two different data sets, i.e. positive and negative, Click on Classify. Classification dialogue box will appear, click on ‘Exclusion’ and enter ‘<0’ or ‘>0’ depending on requirement. Click on ‘OK’ to analyse the included data.
11. Click Ok to Visualise map with change in oil production. For Better Visualization, Go to Layout View and add Legend with Active layers as ‘Legend Items’.
12. To create a Graphical visualisation of Wells vs Change in Oil Production, Go to attribute table of the layer, By right clicking on Layer name then ‘Open Attribute Table’, Click on Table options Button then Select ‘Create Graph’.
13. Create Graph Wizard will appear. To visualise both time period in same graph click on ‘Add’, Select ‘Add Series’. For both Series, Select Graph Type, Data Source Layer, Value Field and Data for each axis accordingly. Select Next, Enter details related to Title, Legend and other visual parameters. Click on Finish to generate Graph.
14. To create a Contour visualisation, Kriging’, of Wells vs production, Go to Toolboxes in Catalog dialogue box followed by ‘System Toolboxes’ > ‘3D Analyst Tools.tbs’ > ‘Raster Interpolation’ > ‘Kriging’ OR search Kriging from ‘Search’ Dialogue box by pressing ‘Ctrl+F’.
15. Kriging Dialogue box will appear. Select ‘Input point features’ as Parent layer and ‘Z value field’ as change in oil production. Enter path for Output surface raster, Select Kriging method and Semivariogram model according to requirements. Click ‘OK’ to generate layer with Kriging result. Insert legend for better understandability of map and plotted data

XIII. Conclusion

Exploratory Data Analysis of UK oil production data was done for period 1980 to 2010. Different geographical attributes were analysed along with oil production. Relation between Geo-spatial reference and oil production was found and concluded. Other applications and pattern detections were done and suggested for the UK oil production data.