|  | **OIL/GAS EXPLORATION AND PRODUCTION – PHASE 1**      ORANGE HOMEWORK TEAM 8  Karthick Krishna Balaji  Camille Carter  Margeaux Johnson  Dillard McMichael  Nish Torane    February 07, 2023 |
| --- | --- |

**Table of Content**

[**Overview**](#_heading=h.17dp8vu) **1**

[**Methodology and Analysis**](#_heading=h.3rdcrjn) **1**

[**Data Used**](#_heading=h.haldq4efravl) **1**

[**Normality Check**](#_heading=h.ab8240tk7x6x) **1**

[**Kernel Density Estimation**](#_heading=h.g5cmkkuo4b3l) **2**

[**Estimating For 2023**](#_heading=h.8llr5zil0q9q) **3**

[**Results and Recommendations**](#_heading=h.cq2321jxrcvd) **5**

[**Recommendations**](#_heading=h.vns6zifpn2tr) **5**

[**Conclusion**](#_heading=h.esxwwxo6toq1) **5**

[**Appendix**](#_heading=h.exj2x5eo9xv8) **5**

**Oil/Gas Exploration and Production – Phase 1**

# **Overview**

Compagnie Pétrolière et Gazière (hereafter the “Company”) has tasked Orange Homework Team 8 (hereafter the “Analysts”) to simulate future values of 2023 drilling costs. The Company has provided two datasets containing historical drilling costs from 1960 to 2007 and projected costs from 2023 to 2050.

Assuming an average cost applied to oil, gas, and dry wells, the Analysts used trends in recent cost changes provided by the U.S. Energy Information Association to simulate possible 2023 drilling costs. After performing simulations under both the assumption of normality and a Kernel Density Estimate (KDE) of arithmetic changes from 2006-2012, the Analysts found that the KDE resulted in a greater spread of projected cost values. For this reason, the Analysts recommend using a KDE distribution for simulating future cost values, as the broader range gives the Company a clearer picture of all possible outcomes.

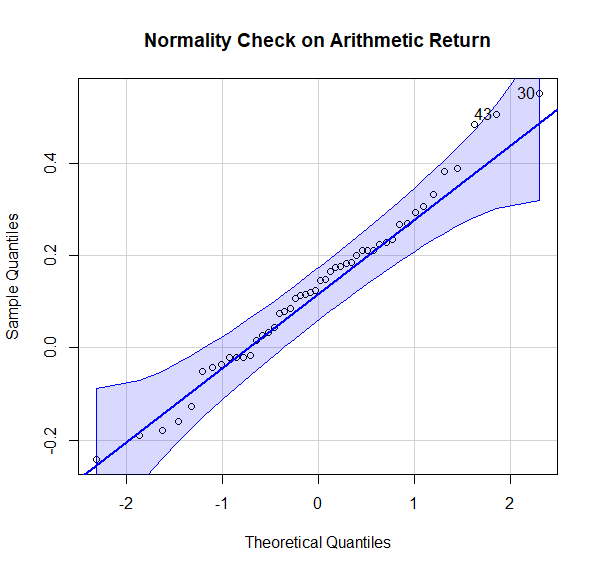
# **Methodology and Analysis**

## Data Used

The Company provided the Analysts with two datasets. The first dataset contains historical Drilling Cost data from 1960 to 2007. The second dataset contains the projected price of drilling costs from 2023 to 2050. For this phase, the Analysts used the first dataset. This dataset includes the cost of drilling crude oil, natural gas, and dry wells in thousands of dollars. The dataset also provided arithmetic returns for each year and each drilling method. As suggested by the Company, the Analysts only used data between 1990 and 2006. The Analysts converted the character variables to numeric and calculated the average annual cost and return as two new variables.

## Normality Check

The Analysts performed a normality test on the arithmetic changes of return in the subset data using a Q-Q plot and the Shapiro-Wilk test. **Figure 1** shows the output from the Q-Q plot, visually representing how the data follows a normal distribution.

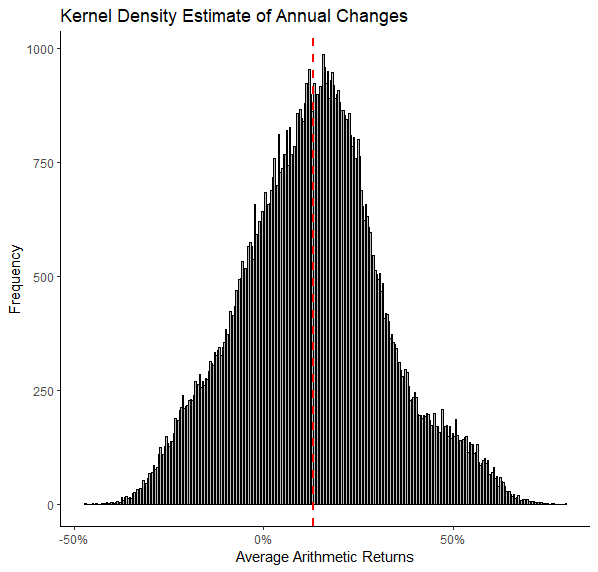


**Figure 1: Normality Check on Arithmetic Return**

Confirming the arithmetic changes follow a normal distribution, the Analysts performed the Shapiro-Wilk test on the same data and obtained a p-value greater than the significance level of 0.05.

## Kernel Density Estimation

The Analysts performed the KDE on the annual arithmetic changes of return in the subset data. After first identifying the bandwidth, the Analysts used 100,000 sample points to estimate the kernel density. The graphical output can be seen in **Figure 2**.



**Figure 2: Kernel Density Estimate of Annual Changes**

The red dotted line represents the mean value (13.13%) of the average changes, as seen in **Figure 2**. The minimum and maximum values of average arithmetic returns were between -47.05% and 79.53%, respectively.

## Estimating For 2023

The Analysts ran two simulations to calculate future values of 2023 drilling costs. The first method uses a normal distribution for the 2006-2012 data, and the second uses KDE for the 2006-2012 data. A report details the changes in cost for the 2012-2015 data as decreasing on average by 9.17% per year with a minimum of 7% and a maximum of 22%. The changes in cost for the 2015-2023 data increase on average by 5% per year, with a minimum of 2% and a maximum of 6%. Given this information, the Analysts used two triangular distributions in both simulations to estimate the remaining years: one for the 2012-2015 data and another for the 2015-2023 data, given the previous information. To simulate possible future values of 2023 drilling costs, the Analysts also used the average drilling cost in 2006 as the initial cost and ran 100,000 simulations.

The first simulation to identify the 2023 drilling cost involves the product of Initial cost, KDE for 2006-2012, and the triangular distributions for 2012-2015 and 2015-2023. The resulting distribution is pictured in **Figure 3** in the Appendix. The red line from the above figure shows the initial cost from 2006.

The second simulation to identify the 2023 drilling cost involves a product of Initial cost, normal distribution for 2006-2012, and the triangular distributions for 2012-2015 and 2015-2023. The resulting distribution is pictured in **Figure 4** in the Appendix. The red line from the above figure shows the initial cost from 2006.

# **Results and Recommendations**

For result:

|  | Minimum | 1st Quartile | Median | Mean | 3rd Quartile | Maximum |
| --- | --- | --- | --- | --- | --- | --- |
| KDE | 587067 | 3095166 | 4137836 | 4472709 | 5477646 | 21168103 |
| Normal | 547877 | 3162403 | 4182659 | 4466106 | 5454065 | 18646863 |

Table: Summary Statistics for Simulated 2023 Drilling Costs

Initial cost: 2279800

## Recommendations

The Analysts recommend the Company use the KDE distribution to estimate the arithmetic changes for the years 2006 - 2012. Compared to the normal distribution estimation, the stimulated costs of the KDE distribution had a broader spread that better accounted for the possible cost projections. This information for 2023 costs will help best anticipate and prepare for future operations. Pivoting to assuming a KDE distribution will improve the Company’s drilling cost estimation going forward.

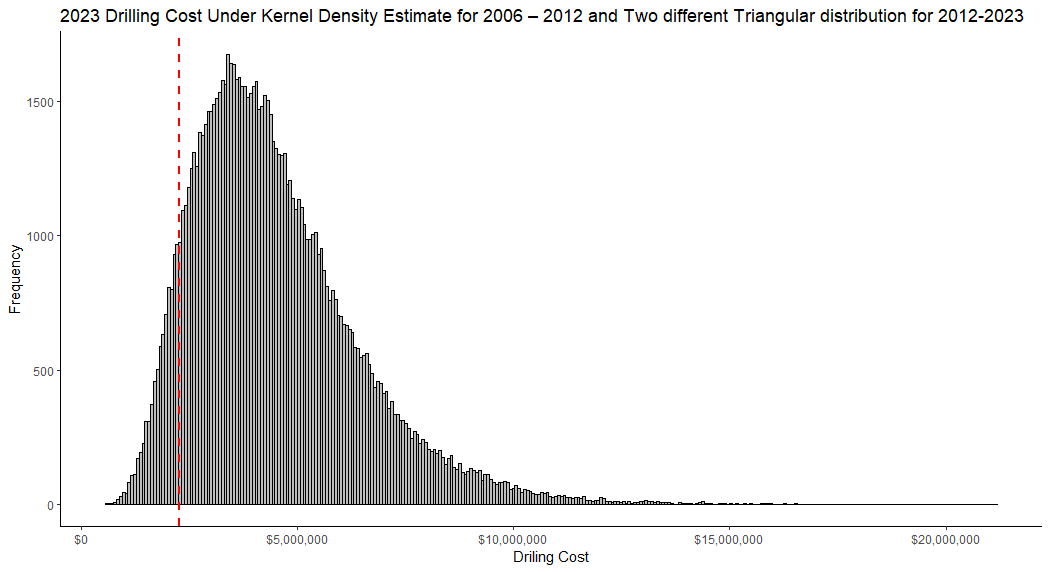
In addition, the Analysts recommend that the Company simulate the average costs for oil, gas, and dry wells individually. Since the aggregated simulation was positively skewed, it would be helpful to understand what drives those inflated costs. Conducting individual simulations for each type and analyzing their distributions will allow a better understanding of which product type is associated with the most risk. The Company would then be able to better allocate its resources and account for possible risk.

# **Conclusion**

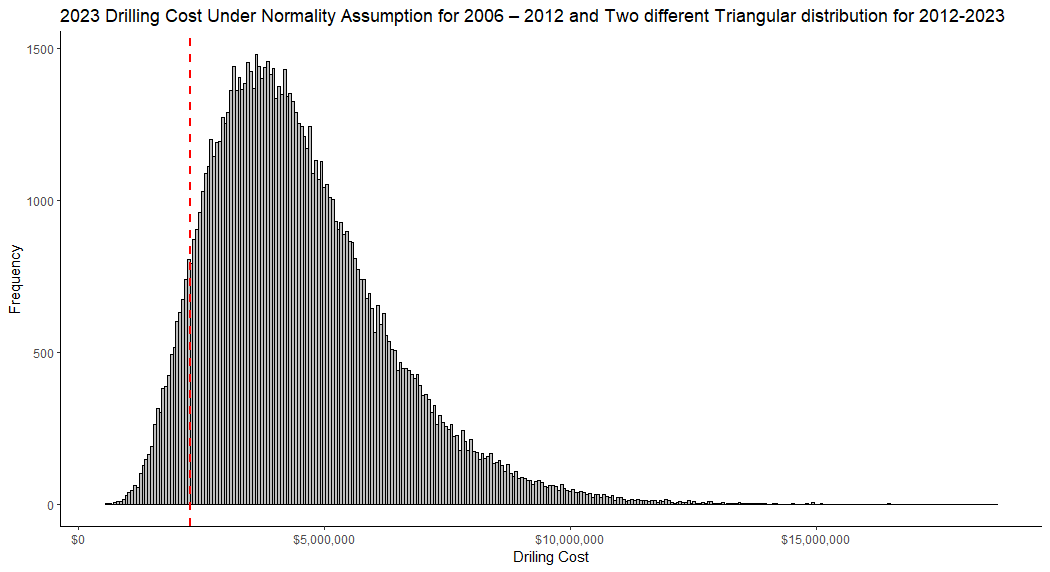
This report details the process of simulating possible future values of 2023 drilling costs using both normal distribution estimation and kernel density estimation. The Analysts only used the cost data from 1990-2006 due to tremendous changes in the cost structure in the previous decades. From 2007 to 2023, the analysts used two different approaches - one that combined a normal distribution estimation with two triangular distribution estimations and the other that combined a kernel density estimation with two triangular distribution estimations. The Analysts also checked and verified that the cost changes from 2007-2012 satisfied normality. The initial cost used to model the 2023 drilling cost came from the 2006 price estimate which was $2,279800.

The 2023 drilling cost results obtained from both the approaches showed that the one that included the KDE estimation had a wider distribution variance than the one with normal distribution. The mean drilling cost with the KDE estimation was also higher than the one with normal distribution estimation by $6603. Based on the results, the Analysts recommend the Company to use the approach that includes the KDE distribution. This will ensure the most comprehensive estimation of drilling cost based on a wider distribution variance. It would also enable the Company to better account for risk and help with resource allocation.

# **Appendix**



**Figure 3: 2023 Drilling Cost under Kernel Density Estimate for 2006-2012**



**Figure 4: 2023 Drilling Cost under Normality Assumption for 2006-2012**

**Homework Report Checklist**

The team member(s) responsible for checking each item should enter their initials in the field next to each question. All items should be addressed before submitting the assignment with the initial checklist attached.

**Sections & Structure**

**Overview**

|  | Is the overview concise? |
| --- | --- |
|  | Does it provide context about the business problem? <Content> |
|  | Does it briefly address your team’s work, quantifiable results, and recommendations? <Action> |
|  | Does it offer audience-centered reasons for recommendations? <Context> |

**Body Sections**

|  | Does the report body include information on methods, analysis, quantifiable results, and  recommendations? |
| --- | --- |
|  | Is content grouped into appropriate sections (methodology, analysis, results, recommendations)? |

**Conclusion**

|  | Does the report have a conclusion? |
| --- | --- |
|  | Does the conclusion sum up the report and emphasize relevant takeaways? |

**Structure**

|  | Does each major section have a heading? |
| --- | --- |
|  | Are sections, subsections, and paragraphs organized logically for easy navigation? |

# 

**Visuals**

**Introduction, Discussion, and Captions**

|  | Is each visual introduced in the text before it appears? |
| --- | --- |
|  | Is each visual close to where it is introduced? |
|  | Does each visual include a title with the following information: type (table or figure), number, and a descriptive caption? |
|  | Is each visual discussed and interpreted in the text? |
|  | Are figures and tables numbered separately? |
|  | Are table captions above the table? Are figure captions below the figure? |

**Visual Design**

|  | Do figures/tables use audience-friendly labels rather than variable names? |
| --- | --- |
|  | Are the visuals easy to interpret? |
|  | Are the visuals appropriately sized? |
|  | Do tables appear on one page (not split between 2 pages)? |
|  | Are legends and axis labels included for figures? |
|  | Are numbers in tables right aligned? |
|  | Are the visuals designed well (ex: re-created in Word or Excel, not blurry or stretched,…)? |

**Document Design**

**Title Page Design**

|  | Does it include a descriptive title? |
| --- | --- |
|  | Does it state the team name, team members’ names, and the submission date? |

**Table of Contents Design**

|  | Does it list all the major sections of the report with corresponding page numbers? |
| --- | --- |
|  | Do the page numbers and sections in the Table of Contents match the report? |

**Document Design for Entire Report**

|  | Is a standard typeface (Calibri, Arial, etc.) used? |
| --- | --- |
|  | Is the size of the body text between 10-12 pt.? |
|  | Are headings and subheadings used to organize information? |
|  | Are distinctive text styles (bold, italic, etc.) used to distinguish between heading levels? |
|  | Are text styles for headings used consistently (ex: all level-one headings are bold)? |
|  | Are all paragraphs an appropriate length (fewer than 12 lines)? |
|  | Is white space used to indicate paragraph breaks? |
|  | Are bullet lists used for a series of items and numbered lists to show a hierarchy? |

# 

**Writing Style and Mechanics**

**Spelling and Capitalization**

|  | Are spelling errors located and corrected? |
| --- | --- |
|  | Is spelling consistent throughout (no switching between acceptable spellings)? |
|  | Is capitalization used appropriately (proper nouns, etc.)? |
|  | Is capitalization of words consistent throughout the report? |

**Grammar and Punctuation**

|  | Are verb tenses used appropriately? |
| --- | --- |
|  | Are marks of punctuation used appropriately? |
|  | Is subject-verb agreement used in every sentence? |
|  | Is the grammar checker updated and are underlined grammar issues addressed? |

**Writing Style**

|  | Are all sentences in the report easy for your audience to understand quickly? |
| --- | --- |
|  | Are most sentences written in active voice? |
|  | Are idioms and vague words eliminated from the report? |
|  | Are acronyms introduced before being used? |
|  | Are well-written topic sentences included at the beginning of each paragraph? |
|  | Are lists parallel? |
|  | Is the appropriate point of view used when addressing your audience or describing team actions? |