

1-1-2018

Formulación de un modelo lineal de regresión múltiple para determinar el personal necesario en los proyectos de la empresa DR&MR

David Alexander Murillo Rojas
Universidad de La Salle, Bogotá

Follow this and additional works at: https://ciencia.lasalle.edu.co/ing_industrial

Citación recomendada

Murillo Rojas, D. A. (2018). Formulación de un modelo lineal de regresión múltiple para determinar el personal necesario en los proyectos de la empresa DR&MR. Retrieved from https://ciencia.lasalle.edu.co/ing_industrial/91

This Trabajo de grado - Pregrado is brought to you for free and open access by the Facultad de Ingeniería at Ciencia Unisalle. It has been accepted for inclusion in Ingeniería Industrial by an authorized administrator of Ciencia Unisalle. For more information, please contact ciencia@lasalle.edu.co.

**FORMULACIÓN DE UN MODELO LINEAL DE REGRESIÓN MÚLTIPLE PARA DETERMINAR EL
PERSONAL NECESARIO EN LOS PROYECTOS DE LA EMPRESA DR&MR**

David Alexander Murillo Rojas

**UNIVERSIDAD DE LA SALLE
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA INDUSTRIAL
2018**

**FORMULACIÓN DE UN MODELO LINEAL DE REGRESIÓN MÚLTIPLE PARA DETERMINAR EL
PRESONAL NECESARIO EN LOS PROYECTOS DE LA EMPRESA DR&MR**

David Alexander Murillo Rojas

Trabajo de Grado para Optar por el Título de:
Ingeniero Industrial

Director (a)
CARLOS ANDRÉS LONDOÑO ARANGO
Ingeniero Mecánico
Máster Ingeniería Industrial

UNIVERSIDAD DE LA SALLE
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA INDUSTRIAL
2018

DEDICATORIA

A mi mamá que soportó y me ayudó en cada paso de esta etapa

A mi familia, que están ahí siempre que los necesito.

Y a Dios que me permitió poder llegar hasta acá con mucho esfuerzo y su ángel.

David Alexander Murillo Rojas

AGRADECIMIENTOS

A DR & MR Comunicaciones y Servicios LTDA. Que me permitió realizar esta investigación y usar sus datos, que me apoyo en todo este camino y me brindó la oportunidad de obtener experiencia a través de él.

A los profesores Alex Pinzón y Carlos Arango, por la asesoría, dirección apoyo y paciencia durante la realización de este proyecto; por enseñarme a ver los temas desde otra perspectiva y que es necesario esforzarse para alcanzar cada meta propuesta.

A la Universidad de La Salle que a lo largo de toda esta etapa me brindo las herramientas y apoyos necesarios para lograr el objetivo, que tuvo la paciencia y dedicación para sacar este nuevo proyecto adelante.

RESUMEN

DR & MR Comunicaciones y Servicios LTDA. es una organización enfocada a realizar proyectos que tienen como principal objetivo la instalación de fibra óptica y sus equipos correspondientes; la empresa está encargada desde el diseño, planeación y presupuesto hasta la construcción de cada proyecto; los cuales se desarrollan en ciudades fuera de Bogotá; Causando así que la empresa deba estar tener pendiente sobrecostos en la operación por traslados y viáticos que necesita cada técnico para el desarrollo de un proyecto.

Partiendo de los retos logísticos y sobrecostos previamente mencionados el siguiente documento plantea un modelo que permite a la empresa DR&MR definir la asignación de recursos a usar en cada uno de sus proyectos, mientras maximiza las ganancias de esta, para ello, toma como base el tiempo en el que requiere terminar cada uno de los proyectos.

Para realizar el proyecto, inicialmente se recopilaron datos de utilidad sobre la empresa en una matriz de gerencia, como lo son mano de obra utilizada definida por perfil, tamaño y tiempo de cada proyecto, incluyendo tiempos desaprovechados en los cuales no se pudo realizar ningún tipo de actividad en pro al desarrollo del proyecto, bien sea por retrasos causados por la logística con el cliente o temas ambientales en el terreno donde se está trabajando. Posteriormente, se construyó una base de datos con 10 variables, a partir de las cuales se realiza un análisis inicial que permite determinar la influencia de cada una de la variable Tiempo; para esto, se diseñó una matriz de correlaciones y unas gráficas de dispersión que permitieron la depuración de variables para cada uno de los ítems a controlar. Una vez descartadas aquellas variables independientes que no cumplían con el requisito estadístico de guardar una relación lineal con la variable dependiente y además no contaban con una correlación de (0.6), se realizaron un conjunto de modelos, creando una tabla en donde se evidencia el comportamiento de cada uno de ellos según las variables utilizadas y las observaciones depuradas; dichas observaciones por cuestiones externas,

contaban con una desviación residual mayor a 2 y menor a -2, claro indicio de que estas sesgaban el resultado del modelo planteado, razón por la cual durante el proceso de creación de los modelos se depuraban las observación que contenían más ruido en la ecuación.

Por último y con el modelo obtenido se realizó una prueba con una muestra test de 5 proyectos para probar la capacidad predictora del modelo, posteriormente y teniendo en cuenta los resultados obtenidos se hizo una pequeña simulación del ROI del año 2017 si la empresa hubiera aplicado el modelo, de esta manera comparar los cambios que tendría la facturación de la empresa al emplear el modelo.

Finalmente se montó una tabla en Excel con todas las variables del modelo, que permitió definir la fuerza laboral necesaria para cumplir con los tiempos exigidos por el cliente.

Palabras Clave: Regresión Lineal Múltiple, Correlaciones, Estandarización, Caracterización de Procesos.

ABSTRACT

DR & MR Communications and Services LTDA. is an organization focused on carrying out projects whose main objective is the installation of optical fiber and its corresponding equipment; the company is in charge from the design, planning and budget until the construction of each project; Those who are in cities outside of Bogotá; Causing so that the company should be having outstanding costs in the operation by transfers and per diem that each technician needs for the development of a project.

Part of the logistical challenges and cost overruns are shown in the following document, instead of in the companies. DR & MR defines the allocation of resources to use in each of its projects, while maximizing the profits of this, based on the time in which each of the projects is required.

To carry out the project, to do this, collect useful data about the company in a management matrix, such as labor. of activity in favor of the development of the project, either by delays due to logistics with the client or environmental issues in the field where he is working. Subsequently, a database with 10 variables was constructed, from which an initial analysis is carried out to determine the influence of each of the variables Time; for this, a matrix of correlations and dispersion graphs was designed that allowed the purification of variables for each one of the elements to be controlled. Once discarded variable variables that were not met with the statistical statistic a linear relationship with the dependent variable and that did not have a correlation of (0.6), it is a set of models, a table is created where the behavior is evidenced of each of them according to the variables used and the refined observations; These are the observations that are presented, the continuous responses and the residual deviation greater than 2 and less than -2, the result of the proposed model, the reason why during the process of creation of the models the observations that contain More noise in the equation.

Finally, with the model, a test was obtained with a sample, a test of 5 projects to test the prediction capacity of the model, and after considering the results, a small simulation of the ROI of the year 2017 was made. Model, in this way.

Finally, a table was assembled in Excel with all the variables of the model, which defines the workforce necessary to meet the times demanded by the client.

Key Words: Multiple Linear Regression, Correlations, Standardization, Process Characterization.

Tabla de Contenido

1. INTRODUCCIÓN	14
2. Contexto	16
2.1. Identificación del Problema.	16
2.2. Justificación del Proyecto	19
2.3. Metodología	22
2.4. OBJETIVOS	24
2.4.1. General	24
2.4.2. Específicos	24
3. Selección del Modelo	25
3.1. Regresión Lineal	25
3.2. Correlación	26
3.3. Estandarización	28
3.7. Regresión Lineal Múltiple	30
4. Definición de Variables	35
4.1. Variables y Parámetros	35
4.2. Actores Involucrados en el Desarrollo del Proyecto	37
4.2.1. Cliente	37
4.2.2. Entorno	37
4.2.3. Empresa	37
4.3. Variables determinadas por el Cliente.	38
4.3.1. Cantidad de Metros.	38
4.3.2. Punto de Partida.	38
4.3.3. Tipo de Cable	39
4.4. Variables determinadas por el entorno	39
4.4.1. Tipo de Apoyo.	39
4.4.2. Cantidad de Enrutamientos.	39
4.5. Variables determinadas por la empresa	40
4.5.1. Horas Liniero.	40
4.5.2. Horas Ingeniero	40
4.5.3. Horas Empalmador	40
4.5.4. Horas Diseñador.	40
4.5.5. Horas Conductor	41

4.6.	Correlación y Multicolinealidad entre variables	42
5.	Modelo de Regresión Variable Tiempo Total	44
5.1.	Elección del Modelo	44
5.2.	Criterios de evaluación	44
5.2.1.	R Ajustado	44
5.2.2.	Anova F	45
5.2.3.	P value	45
5.2.4.	Gráfica de Residuales	45
5.2.5.	Residuos Estándar	45
5.3.	Modelos Obtenido	46
5.4.	Datos iniciales	48
5.5.	Anova F	49
5.6.	Análisis de Residuales.	49
5.7.	Datos Atípicos	50
6.	Analisis de Resultados.	51
6.1.	Aplicación del Modelo tiempo en la empresa.	52
6.2.	Método de Aplicación del Modelo.	53
7.	CONCLUSIONES	55
8.	RECOMENDACIONES	56
9.	BIBLIOGRAFIA	57

Tabla 1 Porcentaje de Multas DR&MR..... **¡Error! Marcador no definido.**

Tabla 2 Correlación entre Variables

Tabla 3 Datos Generales Modelo Tiempo

Tabla 4 Datos Atípicos Modelo Tiempo

Tabla 5 Evaluación de Modelo frente a proyectos reales

Tabla 6 Pronósticos del Uso del Modelo

Tabla 7 ROI Real vs ROI con Modelo

Tabla 8 Tabla para utilizar el Modelo	53
Figure 1 ROI anual empresa DR & MR.....	17
Figure 2 Gráfico de técnicos en campo sin actividades	20
Figure 3 Gráfica de una relación lineal.....	25
Figure 4 Grafico de Residuales para Tiempo.....	49

1. INTRODUCCIÓN

Hoy en día la estandarización de procesos es tomada como una ventaja competitiva que permite a una organización sobreponerse a los constantes cambios que tiene el mercado en general. No obstante, hay proyectos dentro de las organizaciones que se realizan bajo condiciones que dificultan tener un margen claro y específico respecto al uso de recursos requeridos para el desarrollo de cada proyecto, lo anterior causa que los mismo cuenten con una alta variabilidad lo que los vuelvo complejos de planear. Adicional, la variabilidad antes mencionada lleva a las empresas a dificultades económicas debido a la mal asignación de los recursos y, como consecuencia a dificultades para cumplir con los compromisos pactados por los clientes.

Cada organización asigna sus procesos y recursos según sus necesidades esto se debe, a que en muchos casos hay variables fuera de control que pueden afectar los proyectos de diferentes maneras (Casari, 2012), dichas variables tienen impacto en los tiempos de producción o en la cantidad de recursos usados, lo que tiene influencia directa sobre los costos. En el caso de las pymes (pequeñas y medianas empresas) la variabilidad en el proceso es más alta lo que las lleva a tener desafíos en cuanto a las capacidades y recursos necesarios para el desarrollo de un proyecto.

El siguiente proyecto tiene como propósito crear un modelo regresión lineal múltiple para la empresa DR&MR el cual permita definir la fuerza laboral necesaria para cumplir con los requerimientos de un proyecto; dicho modelo se va a realizar

Teniendo en cuenta los costos y tiempos de trabajo, el presente proyecto parte de una base de datos existente la cual cuenta con 10 variables medidas para 100 proyectos (1000 datos) después, es necesario determinar que variables son relevantes para la construcción del modelo para lo cual se hace un análisis de correlación el cual no solo muestra la independencia de una variable con otra.

En el capítulo 4 se utiliza modelos de regresión lineal múltiple para estimar el tiempo necesario para la realización del proyecto para esto, se toman se diseña una tabla en donde se muestra el resultado de cada uno de los modelos, las variables que lo integran, la cantidad de observaciones utilizadas, su nivel de confianza y una conclusión sobre el modelo obtenido.

Finalmente se crea una macro en Excel que permite al coordinador de procesos de DR & MR designar los recursos humanos y de tiempo necesarios para llevar a cabo los proyectos dentro de lo presupuestado con el cliente para cumplir tanto con sus exigencias y con las necesidades de rentabilidad de la empresa.

2. Contexto

2.1. Identificación del Problema.

DR & MR Comunicaciones y Servicios LTDA es una empresa dedicada al diseño, tendido e instalación de fibra óptica. Fue creada en el 2005 y hace cuatro años trabaja como contratista de UFINET S.A, encargándose de realizar dichas actividades en los departamentos de Norte de Santander, Santander, Tolima, Eje Cafetero, Valle del Cauca y Antioquia. En este momento la empresa cuenta con 45 personas de planta divididos entre 13 administrativos y 32 operativos, además de 4 contratistas que le colaboran esporádicamente.

Para realizar la facturación de un proyecto, la empresa acude a una matriz de gerencia, en la cual el coordinador de cada uno de los proyectos debe colocar la cantidad en horas hombre según los perfiles tomo la realización del proyecto, dicha información permite llevar un control de las actividades por cada una de las personas y de igual manera tener una base para calcular los tiempos que se deberían demorar los técnicos en cada actividad según las cantidades que maneje, así de esta manera se evalúan los impactos en el tiempo que pueden llegar a tener variables incontrolables para la empresa tales como el clima, la infraestructura o la altura de la ciudad en la que se está trabajando.

La matriz de gerencia permite identificar los sobrecostos en los que incurre la empresa, ya sea por pago de viáticos a los técnicos o, bien en multas que la empresa debe pagar por no cumplir con los tiempos exigidos por el cliente en el contrato. Las multas son únicamente validas cuando esos retrasos no están justificados.

Durante los últimos 4 años la empresa ha recopilado toda esa información de cada uno de los proyectos que ha realizado. Con base en información recopilado se ha estimado el retorno sobre la inversión desde el año 2014 hasta el último año fiscal el 2017 dando los siguientes resultados.

Fuente A Elaboración Propia

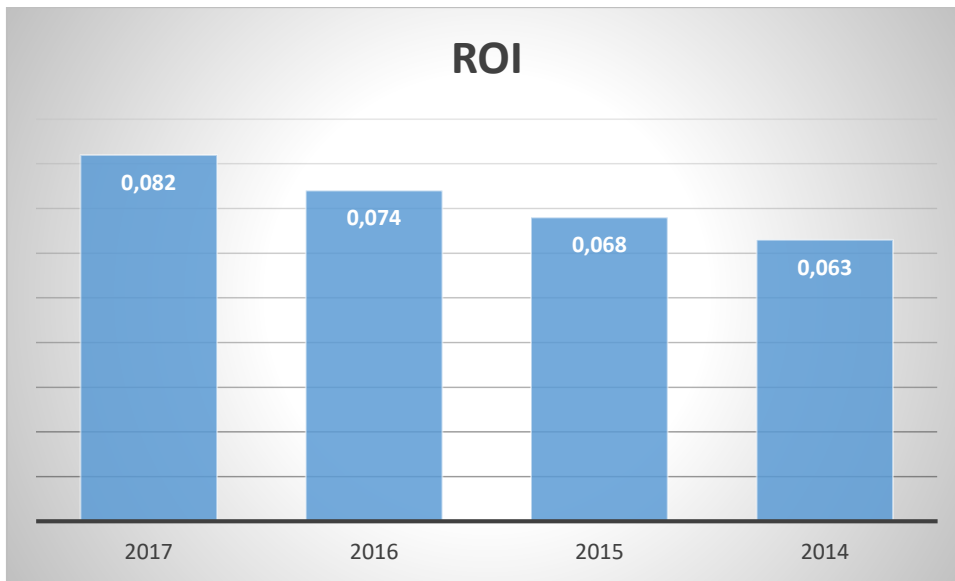
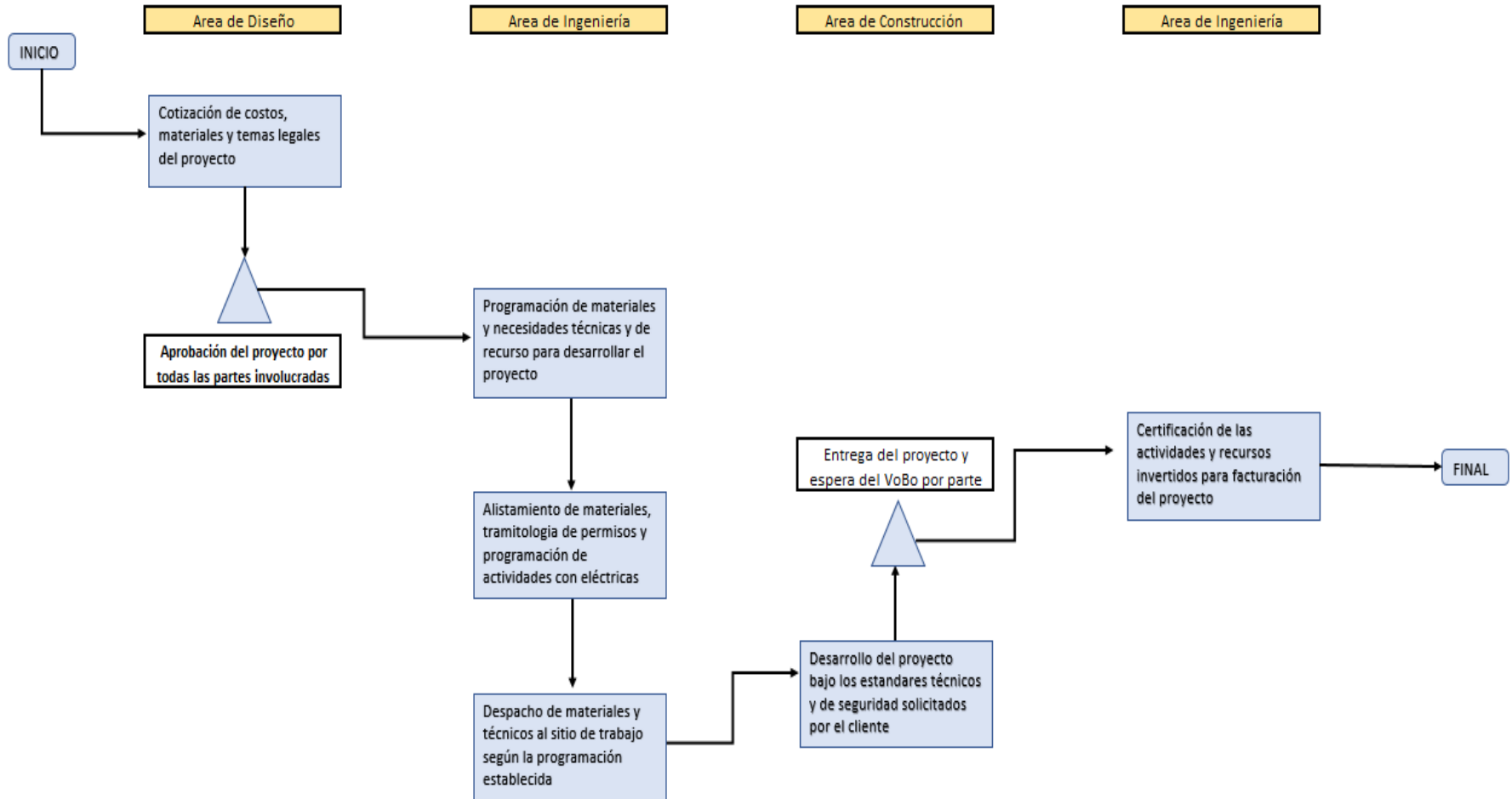


Figure 1 ROI anual empresa DR & MR

Como podemos observar en la gráfica anterior, el ROI ha aumentado cada año, sin embargo, debido a las multas y sobrecostos generados (cuanto es ...referencia), el resultado pudo haber sido mejor. El objetivo principal del proyecto responde a la necesidad de estimar de manera certera los recursos necesarios reduciendo los errores en la estimación y, como consecuencia aumentando el ROI.

Al ser una empresa de servicios de construcción y que gran cantidad del material que se instala lo provee el cliente, el principal recurso de la empresa es la mano de obra, este proyecto busca realizar un modelo que permita definir la mano de obra por perfil, necesaria para cada proyecto, esto con el objetivo de disminuir los y de esta manera aumentar el ROI de la empresa.

Figure 2 Proceso del desarrollo de proyectos empresa DR& MR



Fuente B Elaboración por el autor

2.2. Justificación del Proyecto

Dado que la empresa DR & MR Comunicaciones LTDA. es una empresa que funciona a partir de la realización de proyectos que incluyen actividades como el diseño, montaje e instalación de fibra óptica y equipos que se requieren para su correcto uso, se hace evidente la necesidad de estandarizar procesos que permitan asegurar la calidad y tiempos de ejecución a sus clientes finales y al mismo tiempo permitan a la empresa optimizar su rentabilidad.

Dichos tiempos de trabajo son independientes al tamaño del proyecto ya que los tiempos varían según las necesidades del cliente, por lo que el recurso humano que se debe usar no es necesariamente el mismo para cada uno, lo cual ha llevado a la empresa a incurrir en multas o sobrecostos por mal manejo del recurso humano, dichas multas se hacen efectivas por lo general al no asignar el suficiente recurso para la realización de un proyecto dentro del tiempo estipulado, a continuación, se presentará una tabla que muestra el promedio de multas por proyecto que DR & MR tuvo entre julio del 2017 y enero del 2018.

Tabla 1 Porcentaje de Multas Julio 2017 - enero 2018

Mes	Proyectos	Porcentaje de multas promedio por retrasos en proyectos	Perdida en multas según la facturación	Días promedio de retraso de retraso	Actividad Del Retraso
Julio	32	2%	\$4.640.000	2	Linieria
Agosto	34	2%	\$4.800.000	3	Linieria
Septiembre	38	3%	\$7.350.000	3	Empalmeria
Octubre	30	2%	\$4.600.000	2	Linieria
Noviembre	46	6%	\$15.900.000	4	Linieria
Diciembre	20	2%	\$4.200.000	2	Linieria
Enero	24	4%	\$8.800.000	3	Linieria
Totales	224	21%	\$50.290.000	19	

Fuente C Elaboración Propia

Como se puede observar en la tabla tanto los porcentajes de multas como los días de retraso han aumentado durante los últimos meses estudiados lo cual afecta la facturación de la compañía por las actividades realizadas, también podemos ver que la actividad que genera más retraso es la Liniería (el tendido e instalación de la fibra óptica), razón por la cual en el modelo realizado para calcular el tiempo estimado por proyecto se tomó un enfoque en la variable que representa la cantidad de horas liniero-invertidas en cada proyecto evaluado.

En adición a lo anterior, la empresa también ha incurrido en pago de viáticos a técnicos que en campo están sobrando, o no tienen una actividad a desarrollar que genere algún tipo de ganancia a la empresa, razón por la cual DR&MR ha tomado registro diario en las bitácoras de cada proyecto, en el siguiente gráfico se evalúa cuantos trabajadores están en tiempo muerto, y aun así la empresa debe asumir los viáticos.

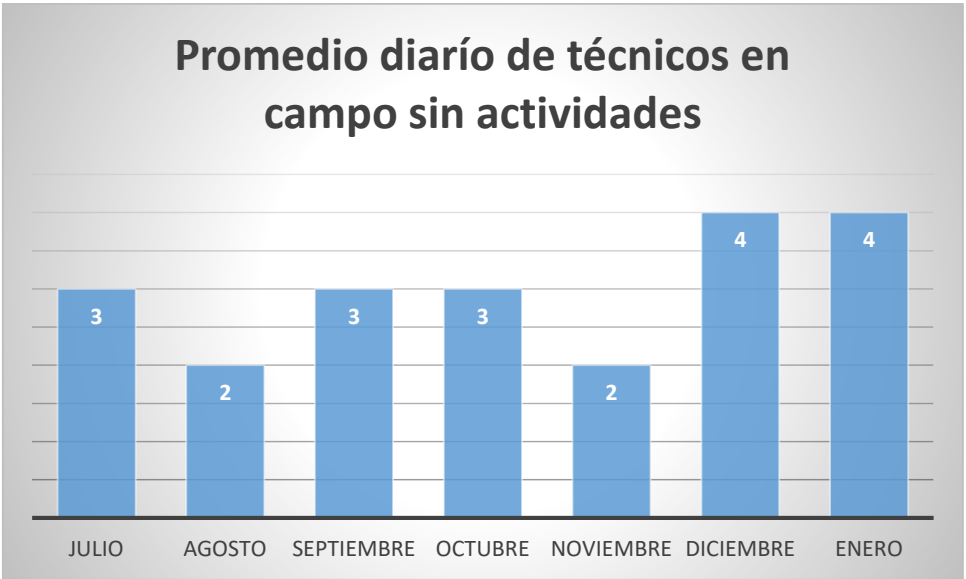


Figure 3 Gráfico de técnicos en campo sin actividades

Fuente: Elaboración Propia

Los gráficos muestran que la empresa adjudica recursos de sobra teniendo un sobrecosto en viáticos y transportes que terminan repercutiendo de manera negativa en las utilidades

que deja cada proyecto a la organización, como se puede observar en el gráfico, son al menos 2 técnicos diarios que la empresa tiene en campo sin realizar ninguna actividad, la empresa paga \$50.000 a cada técnico por los viáticos del día, en los meses de enero y diciembre se están invirtiendo \$200.000 diarios en viáticos de personas que no producen a la empresa, si comparamos esto con las multas nos damos cuenta que el problema de viáticos demás no tiene un efecto financiero alto, sin embargo, afecta la asignación de recursos causando que la empresa incurra en Multas por no tener el personal organizado para cada proyecto.

Por otra parte, es importante analizar la distribución de los fondos de la facturación anual, para identificar el problema real que está teniendo la empresa a nivel económico a causa de los errores a la hora de destinar el recurso humano.

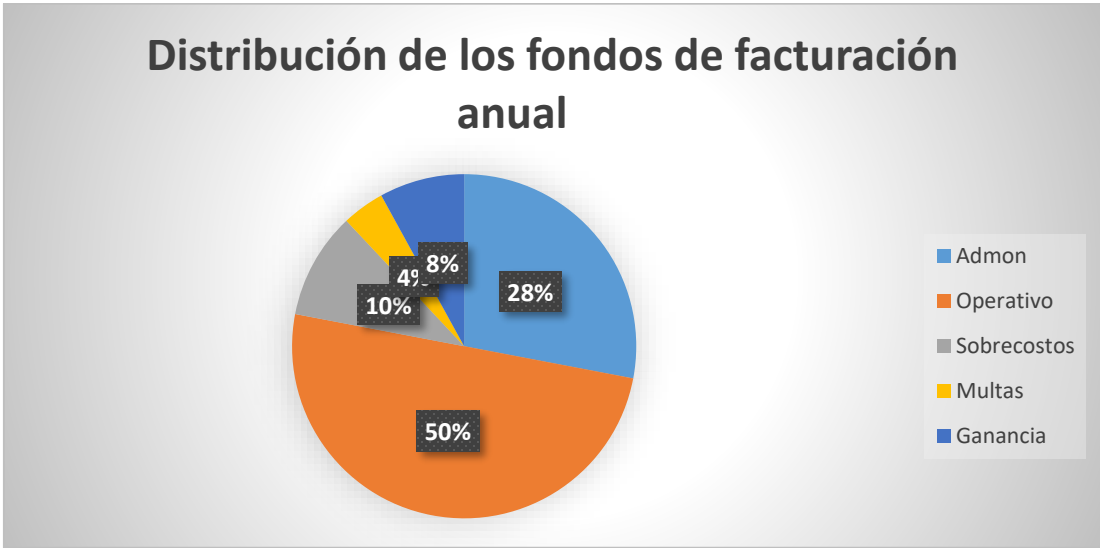


Ilustración 1 Porcentajes de Facturación

Fuente: Elaboración Propia

Como podemos observar, de la facturación, se considera un 14% anual perdido repartido entre multas (4%) y sobrecostos (10%), razón por la cual la ganancia total del año es tan solo el 8%, claro punto de mejora que tiene la empresa, ya que, a pesar de realizar muchos proyectos al año, hay ocasiones donde a la hora de cobrar una factura,

está solo cubre el gasto total del proyecto y no genera ningún tipo de ganancia monetaria a la organización.

2.3. Metodología

Para abordar cada proyecto, DR&MR solicita al cliente las condiciones, restricciones y demás variables que estarán incluidas durante la realización del proyecto, dando como resultado un listado de límites temporales, técnicos y tecnológicos los cuales se tienen en cuenta para, definir el personal necesario para cada una de las actividades del proceso, el dinero que se debe invertir para llevarlo a cabo y una tasa esperada de ganancia de este. Empero a esto, cada uno de estos cálculos sea hace de manera empírica, y únicamente tomando en cuenta las variables expuestas por el cliente, por lo que el incurrir en errores de planeación, que conllevan a gastos extras durante el proceso es un común denominador en la empresa.

Si bien cada uno de los procesos para definir recursos que maneja hoy en día DR & MR son meramente empíricos, ya que la empresa asigna el recurso según un supuesto técnico en donde sin importar las condiciones o variables un técnico debe rendir “x” metros de fibra óptica al día, para controlar esto, hace un año, la empresa lleva registro de cuales han sido las fallas de esto frente al proceso, dando así pautas para tener hoy en día una base en donde se pueda definir no solo las variables expuestas por el cliente, sino aquellas variables implícitas que tiene cada proyecto pero que no se pueden manejar o explicar (Ciudad, clima, permisos, etc.). Y de esta manera ampliar el modelo que se va a realizar, para que este tenga en cuenta la mayor cantidad de aspectos que va a tener el proyecto a realizar.

Ya que las únicas variables donde la empresa tiene control son las horas dedicadas de cada uno de los diferentes perfiles que maneja se realizará una breve explicación de cómo ingresan cada uno en el desarrollo del proyecto y cuál es su influencia en las dos variables que se tienen en cuenta de este.

Recurso Humano de la empresa en el desarrollo de proyectos.

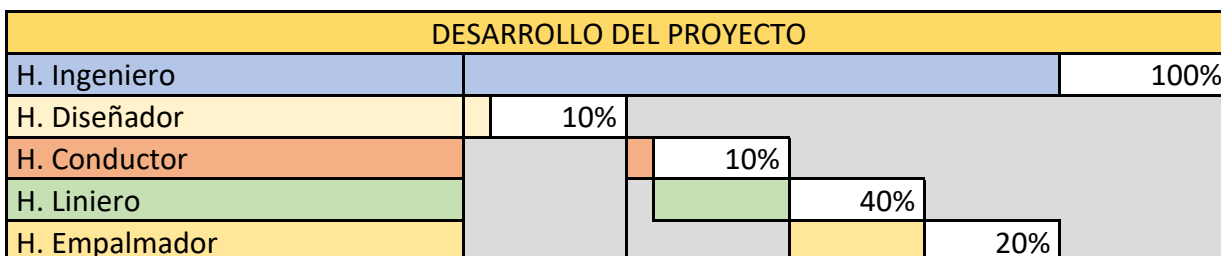


Ilustración 2 Diagrama Gantt Desarrollo de un Proyecto

Como se puede observar el rol del ingeniero acompaña el proyecto en cada una de sus etapas, sin embargo los demás perfiles cuentan con un porcentaje limitado y se tiene como actividad más pesada la Linieria, también hay dos momentos donde el único perfil que interviene es el Ingeniero, el primero es al pasar de diseño a construcción controlando materiales y permisos, y, el segundo es cuando se finaliza la operación, ya que el ingeniero debe realizar unos entregables que permitan cobrar el proyecto.

Fuente D Realización propia

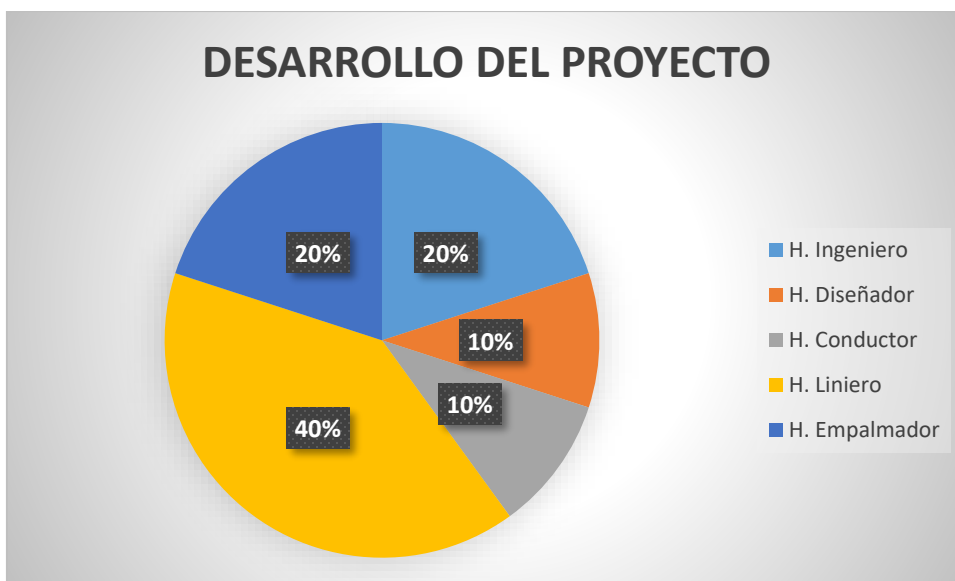


Ilustración 3 Grafico de Desarrollo de un Proyecto.

Con el anterior gráfico nos damos cuenta que la actividad de Linieria es la que más peso tiene en el tiempo de desarrollo de un proyecto, y de igual manera podemos ver como el ingeniero y el Empalmador cumplen también con un rol importante a la hora de reducir el tiempo de realización toma un proyecto, de esta manera tenemos una base que nos da una guía general de que variables deben tener un nivel de relevancia adecuado en el modelo que se está formando.

2.4. OBJETIVOS

2.4.1. General

Modelar la cantidad de horas hombre necesarias por perfil en la empresa DR & MR para la realización de cada proyecto en el tiempo estipulado por el cliente y disminuyendo los sobrecostos de las actividades

2.4.2. Específicos

Definir las variables del proceso usando datos históricos y modelos de correlación para usar las que estén disponibles y sean relevantes.

Modelar usando regresión lineal múltiple para definir la fuerza laboral necesaria para representar el tiempo de cada proyecto según las variables elegidas.

Realizar un macro en Excel que permita utilizar el modelo para realizar la asignación del personal a los nuevos proyectos teniendo en cuenta el tiempo que esté debe tardar en finalizarse

3. Selección del Modelo

3.1. Regresión Lineal

La regresión lineal es una herramienta que permite pronosticar el comportamiento una variable dependiente, frente a cualquier valor que pueda tomar una variable independiente. Partiendo de que estas dos variables cuentan con un supuesto de linealidad; pese a esto, es indispensable realizar un análisis de regresión que provea información sobre el nivel de correlación que tienen las variables que intervendrán en el modelo.

Frente a la Ingeniería Industrial, la regresión lineal puede ser utilizada para crear aplicativos que definan los patrones de comportamiento de una demanda con tendencia (creciente o decreciente), dichos patrones, deben presentar una relación de linealidad entre la demanda y el tiempo, como se muestra a continuación en la siguiente ilustración. (Salazar, 2016)

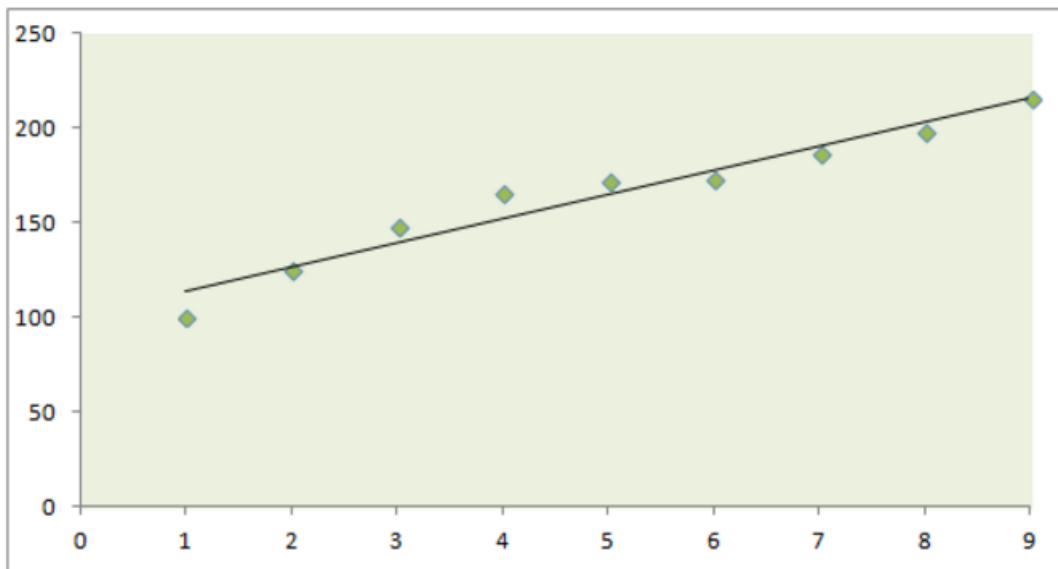


Figure 4 Gráfica de una relación lineal

Fuente: (Lopez, 2006)

Es importante recalcar que la regresión lineal es un modelo básico que únicamente consta de predictor, que sería para él siguiente caso la variable independiente "X", y de una

variable dependiente “Y”, donde cada valor posible de “Y” depende del valor que tome la variable “X” en la ecuación del modelo lineal simple **(1)**:

Ecuación 1 Ecuación de Modelo de Regresión Lineal Simple

$$E(Y|x) = \beta_0 + \beta_1 x$$

Sin embargo, este modelo resulta muy básico a la hora de exponerlo a un caso real y no permite el uso de una cantidad de variables sensata cuando se habla de estandarización de procesos. Ya que al usarlos en modelos más complejos que determinen su comportamiento a partir de dos o más variables independientes se va a tener un resultado sesgado, que más allá de representar una situación va a dar un conjunto de conclusiones erradas, frente a esto el modelo consta de ciertas restricciones como la linealidad que debe presentarse entre las variables. Otra de las falencias que puede llegar a generar este método es que a la hora de graficar debe haber una equidistribución entre residuos, esto para cada par de valores que tome cada una de las variables, para hacer esto en el campo práctico el observador debe hacer uso de su experiencia o heurísticas que permitan definir si es o no aceptable la distribución de residuos en el gráfico del modelo (Mongomery & Runger, 2003).

3.2. Correlación

La correlación es relevante para establecer el grado de dependencia entre las variables estudiadas lo que se convierte en el punto de partida para validar la efectividad en la predicción de los resultados del modelo construido a partir de la regresión lineal. En otras palabras, se puede definir como el objetivo de la correlación “cuantificar la intensidad con la que esas variables se relacionan” (Nuñez, 1992).

Se debe tener en cuenta que la correlación puede existir en diferentes formas. Dependiendo del número de variables puede darse:

-
- Correlación Simple: Dentro se pueden identificar únicamente la variable dependiente y la independiente.
 - Correlación Múltiple: En este caso se estudia la relación de una variable dependiente y dos o más independientes de forma simultánea.
 - Correlación Parcial: Cuando se establece una asociación de tipo múltiple con la diferencia que esta es neta entre dos variables eliminando la influencia estadística de las demás variables independientes.

De acuerdo con la forma que tome la representación gráfica de la función la correlación puede ser rectilínea, parabólica, potencial, exponencial, logarítmica, etc. Por el contrario, si en lugar de los factores anteriores el foco central está en la relación de las variables se puede encontrar la correlación:

- Directa o Positiva: Esta se da cuando un aumento en la variable independiente se traduce en un aumento en la variable dependiente.
- Inversa o Negativa: Se establece cuando al aumentar la variable independiente la dependiente disminuye.

Para determinar la correlación se establece el coeficiente de correlación, el cual indica la asociación existente entre las variables estudiadas, este puede tomar valores desde -1 , denominada Correlación Negativa Perfecta, a $+1$, denominada Correlación Positiva Perfecta donde se situarían dentro de una relación proporcionalmente lineal. No obstante, sin embargo, se tiende a creer que el análisis está sesgado únicamente a una cifra y es importante considerar los siguientes aspectos:

- Dentro de la correlación no se establecen variables explicativas ni explicadas.
- Todas las variables deben poderse cuantificar.

-
- La correlación puede variar dependiendo de los datos que se incluyan en su análisis, es importante incluir la mayor cantidad de datos posibles para que el modelo refleje la realidad.
 - Es relevante dentro del análisis del modelo estudiar las medias y las desviaciones típicas (Moore, 2004).
 - La correlación no indica dependencia (Nuñez, 1992).

3.3. Estandarización.

Conjunto de actividades que tiene como objetivo unificar cada uno de los procedimientos, recursos, tiempo o demás en una organización que utiliza diferentes prácticas para un mismo proceso (Productivity Press Development Team 2002).

Las diferentes etapas que maneja la estandarización de procesos son:

- **Definir el Estándar:** Proceso por el cual se estiman los valores medios donde se unifican cada uno de los ítems (Tiempo, recursos, procedimientos, etc.), esto se puede realizar de varias maneras según el tipo lo que se quiere estandarizar y el modus operandi de la empresa.
- **Informar el Estándar:** Una vez obtenida el estándar, se debe informar mediante un manual de procedimientos, un diagrama de procesos o demás alternativas que permitan transmitir a cada una de las partes involucradas los nuevos estándares ya definidos y que de esta manera se empiecen a involucrar en la operación.
- **Establecer la adhesión de Estándar:** Con cada una de las partes que están involucradas en la estandarización se debe establecer una estrategia que permita implementar la estandarización en la empresa.
- **Mejora Continua:** Como bien se dice, lo que no mejora empeora, cada uno de los procesos que se someten a una estandarización deben estar en control constante para que de esta manera el estándar creado se vaya ajustando a las necesidades de la organización y no se quede en la obsolescencia tan rápido.

3.4. Modelos a base de datos empíricos

Herramienta utilizada para la toma de tiempos principalmente, y funciona a partir de identificar las variables tanto dependiente como independiente del proceso que quiera estandarizar, posteriormente se inicia una recolección de datos, que va desde datos tomados con anterioridad hasta nuevos estudios en donde se verifiquen diferencias o igualdades entre los tiempos de diferentes muestras, para de esta manera definir terminales consistentes (Freivalds, 2004).

Continuando con el desarrollo del modelo cada uno de los datos se deben colocar en una hoja de cálculo y definir mediante la graficación el tipo de relación que hay entre las variables, para finalmente definir un modelo que se acople más a la relación entre las dos partes del procedimiento.

Finalmente, el analista debe probar la formula frente a un estudio de tiempos existentes y analizar cualquier diferencia encontrada, se espera que ninguna supere el 5%, para de esta manera hacer un informe con los resultados y luego realizar el informe con los hallazgos encontrados.

3.5. Formulas Analíticas

Herramienta que se usa para modelar a partir de la información proporcionada por los fabricantes de una maquina o una herramienta, con esta información el analista puede realizar el cálculo de tiempos para las diferentes operaciones que se lleven a cabo con estas herramientas.

3.6. Estudio de Tiempos

El estudio de tiempos es la herramienta más común a la hora de realizar cualquier medición del trabajo y establecer tiempos estándar, este se basa en tres elementos para su propósito (Freivalds, 2004).

- El primero es la estimación de tiempos que debe tomar cada una de las actividades pertenecientes a un proceso, esto con el fin de formar una base preliminar en donde el analista pueda moverse y tener una referencia a la hora de iniciar con la toma de tiempos
- Registros históricos de los estudios de tiempo realizados con anterioridad, en caso de no tenerlos el analista debe iniciar la toma de tiempos basado en la estimación realizada.
- Procedimientos de medición del trabajo, cada una de las actividades debe contar con un protocolo de toma de tiempos acordado entre el analista, el supervisor y el operario

3.7. Regresión Lineal Múltiple.

Teniendo en cuenta que la regresión lineal simple, únicamente contempla una variable predictora, y que esto no es algo que sea sensato frente al caso que nos atañe, se debe ampliar el modelo y utilizar la regresión lineal múltiple.

Si bien en general los principios son los mismos, a la hora de realizar un modelo de regresión lineal múltiple, se deben tener en cuenta la cantidad de variables predictores o independientes que se van a manejar ya que estas, además de manejar el principio de la linealidad, están sujetas a la multicolinealidad, término usado para expresar que dos o más variables independientes se comportan del mismo modo frente al modelo, causando sesgo y que a la hora de realizar la matriz de datos, esta tenga como determinante 0; Frente a los modelos de comportamiento, esto puede dar un sesgo en el modelo o disminuir los efectos reales que tienen los predictores sobre la variable dependiente. Por tanto, cada una de las

variables independientes debe pasar por un conjunto de estudios estadísticos que permitan definir cuál de las variables predictoras cuentan con una multicolinealidad y cuales son aptas para continuar con el modelo.

Si bien cada una de las herramientas usadas en la regresión lineal simple pueden ser usadas para resolver un modelo múltiple, hay que entender que sus métodos de verificación son más complejos, esto debido a que al contar con más de una variable independiente el modelo está más expuesto a errores, por lo que pruebas como bondad y ajuste (que permite definir un R cuadrado, que brinda la confianza que tiene el modelo frente a la realidad, o la matriz de correlaciones), son puntos básicos que permiten darle mayor veracidad al modelo realizado.

A pesar de esto, la naturaleza que maneja la regresión lineal múltiple permite identificar relaciones entre un número mayor de variables, por lo que el modelo cumple con más funciones que estandarizar un proceso de servicios, permitirá determinar las fuerzas laborales relevantes con las que cuenta cualquier empresa dedicada a esta labor:

- Identificar Variables Explicativas: La regresión lineal múltiple le permite generar un modelo que selecciona cada variable que tenga influencia y viendo qué relación tiene con la variable dependiente.
- Detección de Interacciones: Permite observar de qué manera afectan dos variables explicativas por separado y juntas a una misma variable respuesta, dejando claro si la unión de ambas causa la suma, multiplica, resta o divide el efecto que tienen individualmente sobre el modelo.

Como se ha dejado claro, lo primordial antes de realizar cualquier modelo de regresión, ya sea simple o múltiple es un análisis de regresión, que, si bien cambia con cada uno de los patrones modelados, se ajusta a las restricciones y condiciones anteriormente

mencionadas, buscando así que el observador logre un gráfico como el mostrado en la Ilustración 2.

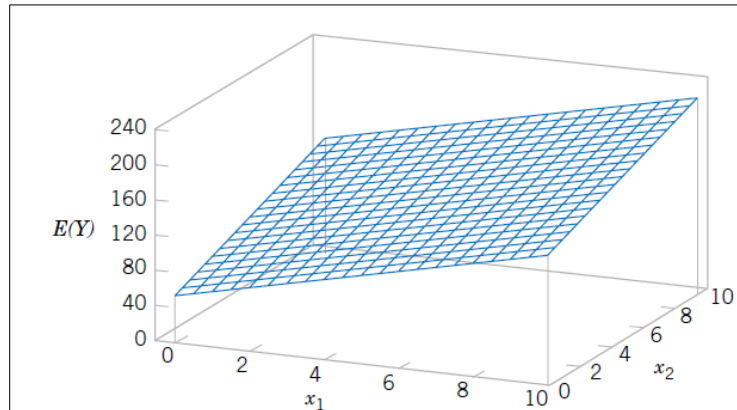


Ilustración 4 Plano para Modelo de Regresión Lineal Múltiple

Fuente: *Applied Statistics and Probability for Engineers*. Wiley.

Sin embargo, cada coeficiente que pueda tener un modelo construido se debe explicar, ya que no es solo un número sino el nivel de intensidad que tiene la relación entre variable independiente y variable dependiente, es por eso que a cada modelo de regresión lineal lo acompaña una tabla de correlaciones, que permite definir la relevancia de cada variable; esto evitará que el modelo tome en cuenta variables que tienen un efecto casi nulo en la respuesta y enfocarse en los que realmente tienen peso.

Finalmente podemos saber de qué manera afecta una variable a otra y esto nos permite plantear escenarios donde veamos cómo reaccionaría la variable “Y” frente a un posible escenario “X”; o que valor se debe usar en “X” para así tener un valor deseado en la variable dependiente “Y”, para tener más claridad en el asunto, el observador busca con cada uno de los estudios estadísticos y con la regresión como tal llegar a un modelo compuesto de la siguiente manera.

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

Ecuación 2 Ecuación de Modelo Lineal Múltiple

En donde “Y” es la variable dependiente, cada “X” es una variable predictora y los β son los coeficientes de peso que tiene cada “X” sobre “Y”; sumado a esto esta ϵ , que es el error que conlleva cada uno de estos modelos y el cual se debe tener en cuenta a la hora de realizar un aplicativo.

3.7.1. Estimación de mínimos cuadrados β

A la hora de analizar un problema de regresión lineal múltiple de forma matricial, como se puede observar en la siguiente imagen, se puede observar de manera más clara las expresiones matemáticas realizadas para hallar los coeficientes de β en un modelo de regresión

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \text{y} \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Figure 5 Matrices de parámetros de Regresión lineal Múltiple

Fuente: (Mongomery & Runger, 2003)

Al observar las matrices el objetivo de la regresión es encontrar los valores para el vector β , pero dichos valores deben ir sujetos a minimizar los valores del vector ϵ ya que lo que se busca es un modelo que represente en su totalidad el comportamiento de la variable

Y, por lo que al realizar la derivada de dicha operación nos encontramos con la siguiente fórmula.

$$X'X\beta = X'Y$$

Ecuación 3 de la derivada del estimador de β

Como el objetivo del ejercicio es calcular los estimados β cada una de las partes de la ecuación se multiplican por $X'X$ dando como resultado la fórmula para la estimación de los mínimos cuadrados de β

$$\beta = (X'X)^{-1} X'Y$$

Ecuación 4 para la estimación de los mínimos cuadrados de β

De esta manera se puede realizar tanto la notación matricial de Y como la de Y' que es el modelo ajustado, esto con el objetivo de obtener un residual o error dado por la diferencia matricial entre (Y) y (Y') .

4. Definición de Variables.

4.1. Variables y Parámetros

Para esta primera instancia se solicitó a la empresa DR&MR el manejo de los datos obtenidos para 100 proyectos, datos que estaban compilados en una matriz de gerencia y en una de facturación Anexo A, esto con el propósito de saber cuáles eran las variables que la organización media de cada uno de los proyectos y apoyados en estas formar una base de datos confiable.

Para empezar a organizar cada una de las variables en la base de datos nueva, se realizó un cuadro conceptual con cada una de las fases con las cuales se debía proceder para de esta manera tener una recolección tanto de datos como de variables más elocuente; Esto ya que cada una de las fases busca descongestionar la gran cantidad de información que se tiene en bruto y ordenarla de manera que se pueda clasificar en grupos diferentes y así mismo la recolección de observaciones de cada una de las variables elegidas sea más sencilla.

Al final la base de datos consta de una serie de variables que, si bien intervienen durante el desarrollo de cualquier proyecto que realice la empresa, están divididas según su naturaleza, el nivel de condicionamiento que le den al proyecto y cuales están a disposición de manejo por parte de DR&MR.

Fuente E Elaboración Propia

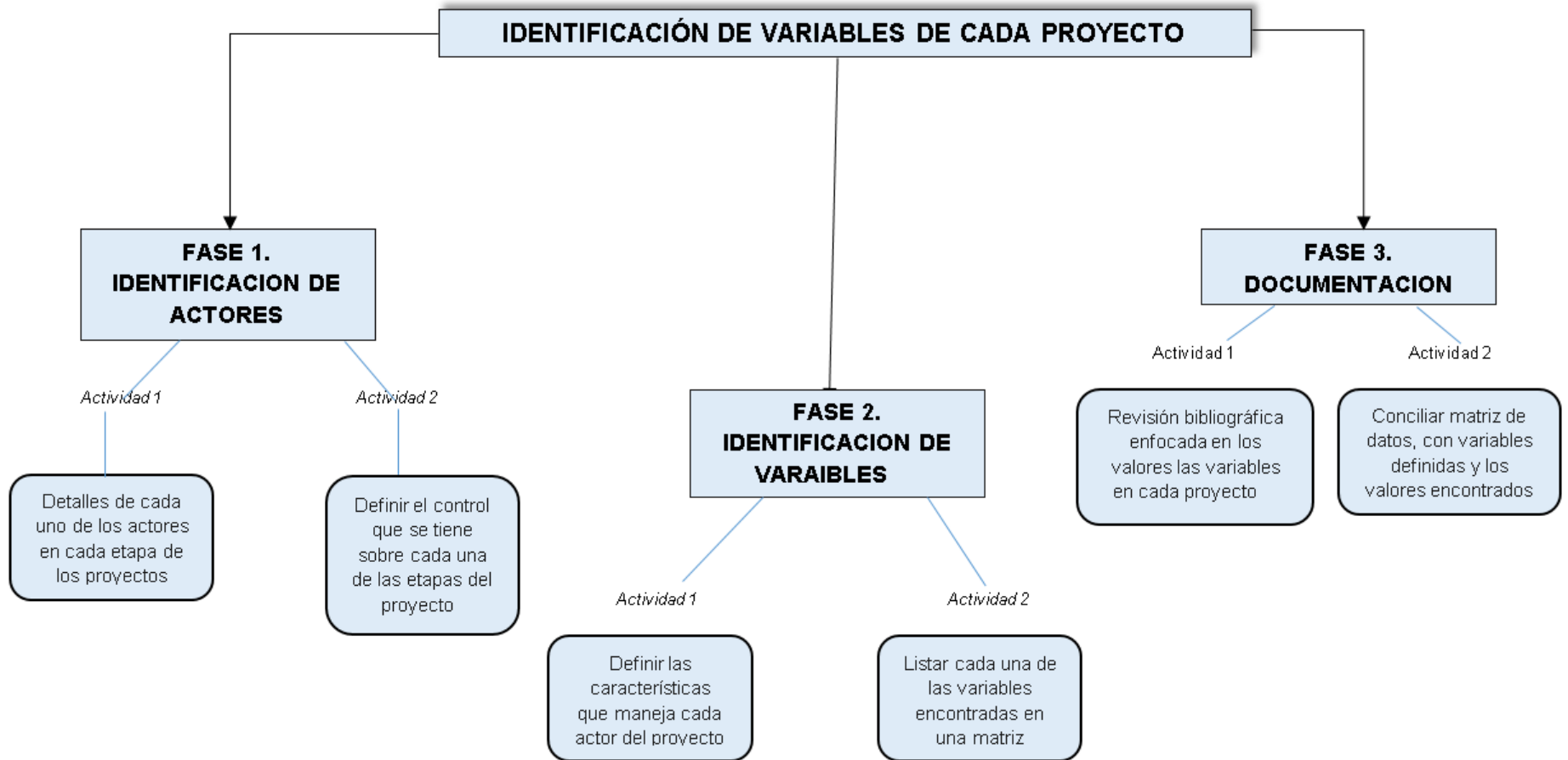


Ilustración 5 Identificación de Variables

4.2. Actores Involucrados en el Desarrollo del Proyecto

4.2.1. Cliente

El cliente es el primer actor identificado y el más importante si en cuestión a variables se habla, esto debido a que él se encarga de asignar una gran cantidad de variables, que además de describir el proyecto como tal, son intocables para la empresa y por lo tanto no están sujetas a ningún tipo de cambio.

4.2.2. Entorno

El entorno es otro de los tópicos importantes para tener en cuenta durante el modelo, esto, debido a que las variables expuestas por este también están fuera de control por parte de la empresa, ya que temas como la ubicación de una última milla (fibra recién instalada) y la cantidad de hilos que se requiere para un proyecto, si bien pueden llegar a ser controlables, se deben seguir ciertos requisitos para crear una red consecuente.

4.2.3. Empresa

Y en el último lugar se encuentra la empresa, ya que debe seguir al pie de la letra el valor de las variables impuestas por el cliente y el entorno a la hora de desarrollar un proyecto, sin embargo, al manejar los recursos necesarios para llevar a cabo esta tiene la necesidad de coordinar de manera que pueda cumplir con los tiempos estipulados con los clientes y aun así tener ganancias que permitan a la empresa continuar y seguir creciendo. Ya que las variables brindadas por la empresa son aquellos recursos manejables que se encuentran alrededor de todo el desarrollo del proyecto, se presenta un Gantt y un diagrama de pastel que permiten ver el momento y el porcentaje en tiempo que maneja cada uno de los perfiles de la empresa en el desarrollo de la mayoría de los proyectos.

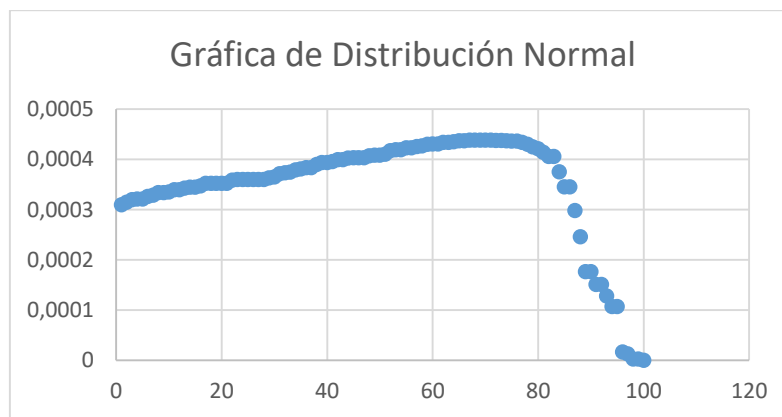
4.3. Variables determinadas por el Cliente.

4.3.1. Cantidad de Metros.

Variable que corresponde a la cantidad en metros de fibra óptica que la empresa debe instalar durante el desarrollo del proyecto, esto con el fin de realizar una conexión entre el cliente asignado y la red existente.

- Es una variable cuantitativa que se mueve entre 113 mts y 5500 mts
- Los datos tienen una media = 871,73 y una Desviación Estándar = 910, 74

Tabla 2 Gráfica D. Normal Variable Cantidad de Metros



Como podemos observar la variable no cuenta con una campana de Gauss que permita definir que dicha variable tiene una distribución normal, por lo que a la hora de incluirla en la regresión, debido a la distribución de sus datos puede aumentar el error y hacer que el modelo pierda predictibilidad.

4.3.2. Punto de Partida.

El punto de partida del proyecto se define en dos categorías, la primera es una Caja de empalme, en donde está ya se encuentra instalada y por lo tanto el técnico únicamente debe realizar la instalación de la fibra óptica dentro de esta, la segunda es una reserva, la cual es un sitio indicado donde el técnico debe acomodar una caja de empalme y posteriormente instalar la fibra óptica correspondiente al cliente, por tal motivo se realizó una variable binaria en donde "0" equivale a caja de empalme

y “1” a reserva de cable, se tomó de esta manera ya que por conocimiento empírico toma más tiempo y por lo tanto mayor beneficio partir de una reserva de fibra óptica, sin embargo dicha información se verificará con las gráficas.

4.3.3. Tipo de Cable

Esta variable expresa la capacidad del cable que se va a instalar en cada proyecto, esta se incluyó debido a que la capacidad correspondiente a 6 hilos es de tan solo 8 mm de diámetro haciendo que su manejo e instalación sea sencillo, mientras que las de 12, 24 y 48 hilos son de 13,2 mm por lo tanto es más difícil su manipulación e instalación en los diferentes tipos de infraestructura, en un principio se va separar cada categoría, sin embargo dados los resultados arrojados por las gráficas se puede simplificar a fibra delgada y ancha.

4.4. Variables determinadas por el entorno

4.4.1. Tipo de Apoyo.

La variable tipo de apoyo se refiere a la infraestructura que la empresa va usar como apoyo para instalar la fibra óptica hasta el cliente, se definen en 2 tipos, postes y cámaras, para poder ingresarlas al diseño se definió una variable binaria de la siguiente manera, 0 para cámaras y 1 para postes, esto se debe a que instalar fibra en un poste conlleva mucho más tiempo y un técnico especializado que cuente con el certificado de trabajo seguro en alturas, mientras que por cámaras la fibra solo va adherida a un muro siendo más sencillo el proceso.

4.4.2. Cantidad de Enrutamientos.

La cantidad de enrutamientos corresponde a el número de veces que el técnico debe realizar la actividad de empalmaría para dejar en alta al cliente, en principio esto depende de lo lejos que esté el cliente de la red principal o la ciudad donde se trabaje.

4.5. Variables determinadas por la empresa

4.5.1. Horas Liniero.

Las horas liniero, corresponde al tiempo invertido por cada uno de los técnicos que están encargados de instalar la F.O en las diferentes infraestructuras de apoyo, estos técnicos deben tener un certificado de trabajo seguro en alturas y depende del apoyo a usar puede tomar más o menos tiempo, el tiempo detallado en la variable más allá de estar ligado en la cantidad de cable a instalar, depende de la cantidad de apoyos que deba intervenir, generalmente entre más metros de fibra se deba instalar más

4.5.2. Horas Ingeniero

Corresponde a la cantidad de horas invertidas por el ingeniero encargado de la coordinación y logística del proyecto, si bien esté no actúa de manera activa en el proyecto, hace un acompañamiento a lo largo de sus diferentes etapas, esto con el objetivo de guiar las actividades de cada una de las partes o resolver los inconvenientes que puedan suceder a lo largo del desarrollo del proyecto.

4.5.3. Horas Empalmador

Corresponden a la cantidad de horas totales que debió invertir el empalmador en instalar, configurar y dar en alta el servicio del cliente en cuestión de cada uno de los proyectos, acciones que se tienen en cuenta debido a que son las más delicadas de todo el proceso y de igual manera son las que más restricciones tanto de tiempo como de dinero tienen.

4.5.4. Horas Diseñador.

Variable que corresponde a la cantidad de horas que le tomo al diseñador realizar el levantamiento de toda la información necesaria para realizar la construcción del proyecto, esta actividad por lo general la realiza un técnico; para el cálculo de estas horas, se tiene en cuenta si el tamaño y el lugar del proyecto permite realizar dos

diseños en el mismo día o si por temas de desplazamiento y tiempo la actividad solo se logra realizar uno.

4.5.5. Horas Conductor

Variable que nos muestra la cantidad de horas que tomo el traslado de la cuadrilla al punto de trabajo, este ítem se tiene en cuenta debido a que la mayoría de los proyectos quedan lo suficientemente lejos de Bogotá como para que la cuadrilla tome un día entero viajando, día en el que no adelanta ningún tipo de actividad correspondiente al proyecto.

4.6. Correlación y Multicolinealidad entre variables

	T. TOTAL	ROI DEL PROYECTO	Q. CABLE	TIPO DE CABLE	TIPO DE APOYO	Q. ENRUTAMIENTOS	PARTIDA	H. DISEÑADOR	H. INGENIERO	H. LINIERO	H. EMPALMADOR	H. CONDUCTOR
T. TOTAL	1											
ROI DEL PROYECTO	0,3318101	1										
Q. CABLE	0,4568901	0,50172665	1									
TIPO DE CABLE	0,3877544	0,322075329	0,70774861	1								
TIPO DE APOYO	0,0922209	-0,091541255	-0,18787	-0,094163	1							
Q. ENRUTAMIENTOS	-0,0087708	0,207891646	0,01085712	-0,121072	0,0065326	1						
PARTIDA	0,2059713	0,135817061	-0,0371175	-0,014874	0,11524936	0,127506567	1					
H. DISEÑADOR	0,3587583	0,377747285	0,85976493	0,589542	-0,181081421	-0,019764964	-0,106095	1				
H. INGENIERO	0,6562556	0,375313977	0,79520323	0,617654	0,013733999	-0,118451783	0,0583846	0,678817533	1			
H. LINIERO	0,416889	0,573240851	0,80296982	0,484366	-0,138906272	-0,017538139	0,0374427	0,651543701	0,681793323	1		
H. EMPALMADOR	0,0463158	0,377671786	0,1137631	0,002922	-0,072486123	0,74878583	0,1891626	0,023065901	-0,014015561	0,081952	1	
H. CONDUCTOR	-0,0142138	-0,010294226	0,15852693	0,1910773	-0,095934208	-0,022807853	-0,184627	0,12512673	0,148899685	0,029694	0,052455489	1

Tabla 3 Correlación entre Variables

- Correlaciones altas
- Correlaciones medias
- Correlaciones multilineales

4.6.1. Análisis de table de correlaciones

Como podemos observar la variable dependiente tiempo, puede tener un posible modelo compuesto de las variables H. ingeniero, H. Liniero, Q. cable y tipo de cable, sin embargo, la variable correspondiente a tipo de cable tiene una correlación muy alta con 4 variables dentro de la base de datos por lo que se cumple el supuesto de multicolinealidad e incluirla en el modelo va a tener impactos negativos frente a la capacidad predictora del modelo; por lo que estás son las variables que se tendrán más en cuanto a la hora de desarrollar el modelo ya que son las de mayor correlación frente a la variable que se quiere predecir. De igual manera ocurre con la variable independiente Horas Diseñador, pero a diferencia del tipo de cable está variable no cuenta con una correlación amplia.

5. Modelo de Regresión Variable Tiempo Total

5.1. Elección del Modelo

De todos los modelos explicados durante el contexto del proyecto, se eligió tomar como herramienta la regresión lineal múltiple esto debido a las siguientes características:

- La RLM (Regresión Lineal Múltiple) permite definir de que manera afecta un conjunto de variables a un todo, mientras que los demás métodos se enfocan en realizar estudios de variable a variable y no a todas como un
- Debido a la cantidad de datos que se maneja en el proyecto, durante la evaluación la RLM nos condiciona a eliminar datos que contengan mucho ruido, en otras palabras, durante la realización del modelo se deben ir sacando datos anormales.
- Tomando como herramienta la RLM podemos definir un conjunto de variables irrelevantes en el proceso, variables que solo hacen peso muerto al modelo pero que a la hora de aplicarlo no generan ningún tipo de confiabilidad al modelo.

5.2. Criterios de evaluación

5.2.1. R Ajustado

Prueba que permite definir qué porcentaje de la varianza de la variable predictiva es explicado por la varianza del modelo, sin embargo, al ser ajustado penaliza la inclusión de variables por lo que sin importar la cantidad de variables que se incluyan en el modelo, el R^2 ajustado va a arrojar datos certeros y para el caso que nos atañe en este proyecto se busca que el modelo cumpla al menos con un 60% de nivel explicativo frente a la variable tiempo.

5.2.2. Anova F

Prueba estadística de análisis de varianza que permite definir si la hipótesis nula, que para este caso es que el modelo cuenta con las variables idóneas para explicar el comportamiento de la variable dependiente, permitiendo así rechazarse o no según el valor F significancia obtenido durante el proceso.

5.2.3. P value

Durante la realización de cada uno de los modelos habrá una depuración de aquellas variables que no cumplen con el P-value menor a 0.05, esto debido a que un mayor valor representaría que la variable no tiene relevancia dentro del modelo.

5.2.4. Gráficas de Residuales

La primera gráfica va mostrar los residuales frente a los resultados obtenidos del modelo, esto con el objetivo de verificar el comportamiento de las variables predictor frente al modelo, ya que esta gráfica permite definir la existencia de multicolinealidad y/o el ruido de los datos obtenidos de la base; el segundo es una gráfica de normalidad y homocedasticidad que permita asegurar el valor predictivo que tiene el modelo según la distribución estadística que tengan los residuos.

5.2.5. Residuos Estándar

Al dividir cada residuo por la raíz cuadrada del valor esperado obtenemos los residuos estándar, los cuales tienen una distribución normal, por lo que podemos identificar si el modelo tiende a aumentar los residuos que hay frente a la realidad, también podemos identificar las observaciones que contienen mucho ruido y excluirlas, así que cada observación con un residuo estándar mayor 2 y menos (-2) será retirada del modelo.

5.3. Modelos Obtenido.

Se realizaron 5 iteraciones en donde se sacaron todas las variables que no tuvieran el P value exigido, de igual manera se sacaron observaciones con residuos estándar fuera de los límites establecidos y se realizó un modelo con la única variable predictora aceptada, sin embargo, al no tener un impacto estadístico mayor, finalmente se decidió por un modelo que tiene 3 variables fuera de los parámetros antes mencionados pero que permiten al modelo alimentarse de más datos a la hora de cumplir el la meta de asignar recursos, para revisar cada una de las iteraciones y que se retiraba en esta Anexo B

Tabla 4 Resultado 1 del Modelo

Coeficientes		P-value
Intercept	15,17733172	0,008831582
Q. CABLE	-0,00229606	0,545940929
TIPO DE CABLE	0,070872172	0,54573094
TIPO DE APOYO	0,953440873	0,70914915
Q. ENRUTAMIENTOS	0,560929022	0,600704931
PARTIDA	3,495189495	0,156980409
H. DISEÑADOR	0,104284406	0,724194009
H. INGENIERO	0,931624893	2,25636E-07
H. LINIERO	0,012748171	0,913040977
H. EMPALMADOR	0,026352543	0,894926779
H. CONDUCTOR	0,251563181	0,284043888

Fuente G Elaboración Propia

Tabla 5 Resultado 5 del Modelo

Intercept	12,01953436	1,06063E-05
H. DISEÑADOR	0,135114504	0,356646181
H. INGENIERO	1,179977634	6,07574E-21
H. CONDUCTOR	-0,165295123	0,243481593
H. LINIERO	-0,076795008	0,219372751

Fuente F Elaboración Propia

Tabla 6 Resultado 4 Modelo

Intercept	11,15074514	0,009987303
Q. CABLE	-0,00063777	0,836451595
TIPO DE CABLE	0,063100663	0,501100129
TIPO DE APOYO	1,109597633	0,538255624
Q. ENRUTAMIENTOS	0,711867066	0,364076962
PARTIDA	-	-
	1,477391551	0,402210476
H. DISEÑADOR	0,129560731	0,579430753
H. INGENIERO	1,090373384	2,49549E-13
H. LINIERO	-	-
	0,034542881	0,668241677
H. EMPALMADOR	0,040645586	0,770617212
H. CONDUCTOR	-	-
	0,256253919	0,121695165

Fuente H Elaboración Propia

Teniendo en cuenta los resultados obtenidos las variables predictoras usadas para el modelo son: H.D (Horas Diseñador), H.I (Horas Ingeniero), H.C (Horas Conductor) y H.L (Horas Liniero) contando así con 4 perfiles distintos de trabajo que permite a la empresa abarcar más información a la hora de usar el modelo.

$$Tiempo T = 12.01 + 0.1351(H.D) + 1.180 (H.I) - 0.1653(H.C) - 0.076(H.L)$$

Ecuación 5 Modelo de regresión lineal para variable predictiva Tiempo Total.

5.4. Datos obtenidos del modelo escogido

Iteración	Variabes	Datos	R ajustado
5	3	85	0,785

Tabla 7 Datos Generales Modelo Tiempo

Los datos iniciales muestran la cantidad de iteraciones que se realizaron hasta llegar a las características deseadas, en cada una de las iteraciones se iban retirando de la matriz principal datos con mucho ruido y un residuo estándar mayor a 2 o -2, finalmente para la última iteración se retiraron cada una de las variables que no cumplían con las mínimas condiciones estadísticas; como resultado para el modelo representativo del tiempo solo se usarán 3 de las 10 variables posibles, y el modelo se realizó sobre una base de muestreo de 85 proyectos.

5.5. Anova F

ANOVA				
	<i>df</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	3	4069,119739	102,1082237	1,94E-27
Residual	81	39,85104815		
Total	84			

Como se puede observar, el valor de la significancia F es menor al mencionado en los criterios por lo que el modelo obtenido cumple con el requisito ya acordado de tener un nivel de significancia menor a 0.05, dando como resultado rechazar la hipótesis nula propuesta durante los criterios de evaluación

5.6. Análisis de Residuales.

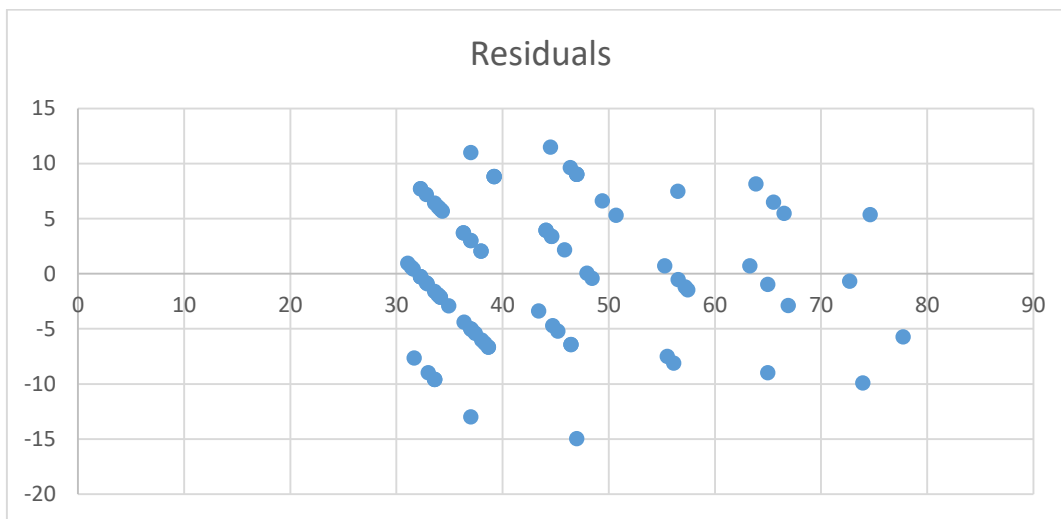
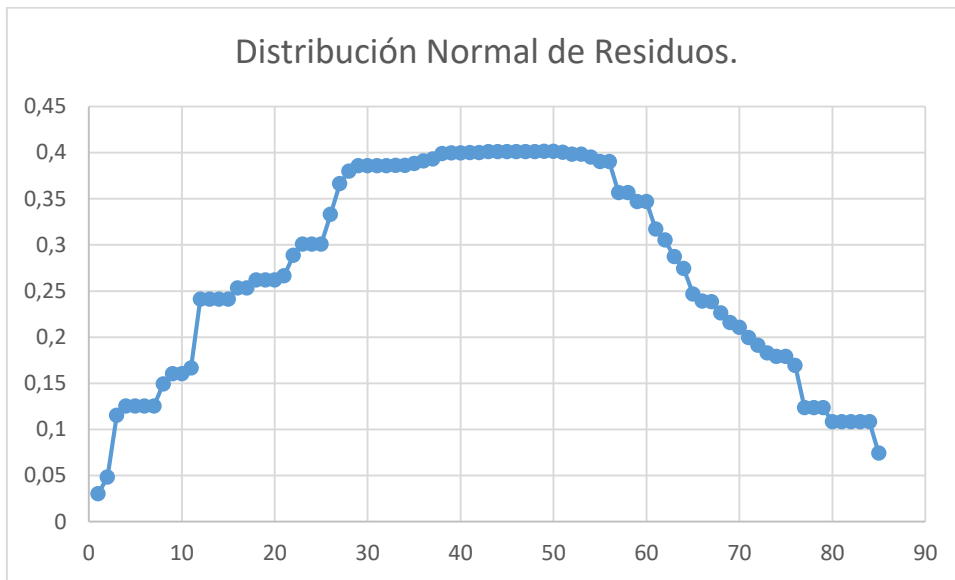


Figure 6 Grafico de Residuales para Tiempo

El gráfico anterior está compuesto por los residuales (Y) y el conjunto de datos obtenido al pasar cada observación por el modelo (X) mostrando una dispersión que, si bien no termina de ser lo lineal que se esperó, da indicios de que el modelo se adecua a la variable que

quiere representar y por tal motivo es un modelo que con aún con el error que tiene debido al ruido de los datos permite trabajar en los objetivos planteados.



Como se puede observar, la gráfica de distribución normal de los residuos obtenidos del modelo, cumple con el requisito de campana de Gauss en donde se logra verificar que los residuos cumplen con una distribución estadística normal con media: 2.40 y Desviación: 0.99

5.7. Datos Atípicos

Tabla 8 Datos Atípicos Modelo Tiempo

		Datos Atípico	
10	46,02191667	-14,02191667	-2,261958544
46	36,6772714	-12,6772714	-2,04504584

Como se observa aún después de realizar varias iteraciones y de minimizar el modelo a una mínima expresión, aún existen datos que generan mucho ruido, sin embargo, estos no se retiraron del modelo final ya que al realizar una iteración más, el R ajustado no tenía un aumento significativo y si se perdían datos en la base del modelo.

6. Analisis de Resultados.

Continuando con la evaluación del modelo obtenido, la nueva fórmula se somete a 5 casos reales tomados de la base de datos, de esta manera podremos ver que tan apegado es a la realidad el modelo y así poder empezar a resaltar la utilidad que tendrá a la hora de asignación de recursos.

Tabla 9 Evaluación de Modelo frente a proyectos reales

DATOS	T. TOTAL	H. DISEÑADOR	H. IN GENIERO	H. CONDUCTOR	H. LINIERO	T. MODELO	DESVIO
101	40	4	24	16	48	35	5
102	48	4	32	4	48	46	0
103	32	4	24	16	48	35	3
104	32	8	20	16	48	30	2
105	64	8	48	12	48	64	0

Como nos muestra la evaluación del modelo, los desvíos obtenidos en la tabla anterior son entre 2 y 5 horas lo cual no representa un día de trabajo, teniendo en cuenta el R^2 es de 0.78, y que la desviación máxima es de 5 horas, usando el modelo la empresa puede garantizar que no sufrirá ningún tipo de retraso no justificado mayor a un día cerca del 80% de los proyectos

6.1. Aplicación del Modelo tiempo en la empresa.

Tomando como base la información dada por la empresa frente a la cantidad de proyectos que tiene y el porcentaje de multas en las que incurre por proyecto, se va a realizar un pronóstico del siguiente mes, sin embargo, esta vez se cambiará la manera de asignar los recursos y se utilizará el modelo para ver el cambio que puede producir en las multas de la empresa.

Tabla 10 Pronósticos del Uso del Modelo

Proyectos	Porcentaje de multas promedio por retrasos en proyectos	Días promedio de retraso de retraso
32	1.04%	1

Tomando como base la información de los últimos 6 meses de la empresa el promedio de cantidad de proyectos realizados al mes es 32, al usar el modelo nos aseguramos de que el 78% de estos proyectos tendrán tan solo medio día de retraso en el peor de los casos, esto llegaría a afectar en 25 proyectos, los 7 proyectos faltantes toman el promedio de multas que viene teniendo la empresa que es 3% por proyecto, dando así el resultado mostrado en la tabla.

Es claro el cambio que puede llegar a representar el uso del modelo para asignar recursos según en los tiempos de entrega de cada proyecto, esto, a su vez afectará directamente la facturación de la empresa disminuyendo las pérdidas causadas al incurrir en las multas y aumentando a su vez el ROI, el cual es el objetivo directo, para evaluar esta hipótesis se va

realizar un cálculo del ROI del año 2017, pero esta vez, adaptando los ingresos totales del año al modelo de asignación de recursos previamente realizado.

Tabla 11 ROI Real vs ROI con Modelo

Ingresos Anuales Reales	Porcentaje de Multas y Sobrecostos	Ingresos reales del año	Margen	ROI
\$2.790.697.674	14%	\$2.400.000.000	0,019583	8,246%

Ingreso Anuales Reales	Porcentaje de Multas y Sobrecostos	Ingresos del Modelo del año	Margen	ROI
\$2.790.697.674	12%	\$2.455.813.953	0,041866	18%

Los cambios observados tanto en el margen, como en el ROI correspondientes al año 2017, en caso de que se hubiera utilizado el modelo, son significativos, con tan solo reducir en 2 puntos el porcentaje que se pierde en multas, el ROI aumenta aproximadamente un 10%.

6.2. Método de Aplicación del Modelo.

Para que el coordinador de personal tenga un método sencillo de utilización del modelo se diseñó una tabla sencilla en Excel que solicite los datos fijos de un proyecto y de esta manera el coordinador pueda empezar a validar el personal relevante para el proyecto, que para el modelo que se obtuvo son los Linieros, a continuación, se muestra una forma básica de utilizar el modelo en Excel.

Tabla 12 Tabla para utilizar el Modelo

TIEMPO REQUERIDO	H. DISEÑADOR	H. INGENIERO	H. CONDUCTOR	H. LINIERO
40	16	26,4	16	29,105947

Como primer ítem tenemos el tiempo, el cual es dado por el cliente como restricción y es de carácter obligatorio para cada proyecto, las horas diseñador y conductor son dadas por el coordinador según la ciudad donde se deba realizar el proyecto, este

es un conocimiento que de antemano debe tener el que manipule la tabla, las horas ingeniero corresponden al 66% de las horas del tiempo requerido, ya que es el promedio de horas de las que hace arte ingeniero de cualquier proyecto, finalmente, la formula calcula las horas liniero necesarias, e esta manera el coordinador define si con uno o dos técnicos de Linierita cumple con dichos tiempos, para el caso que se especifica puede ser un linero durante 3 días o 2 linieros durante 1.5 días.

7. CONCLUSIONES

- La regresión lineal múltiple permite crear y ajustar un modelo, usando las variables recopiladas a lo largo del año y además permite definir las actividades críticas que tiene la empresa durante el desarrollo de cualquier proyecto.
- Si bien el modelo utiliza 4 diferentes perfiles para calcular el tiempo que toma la realización de un proyecto, la Linieria es la actividad crítica y por lo tanto la fuerza laboral que se debe calcular con el modelo, esto debido a que es una actividad a la cual se le puede asignar más de un técnico.
- Los gráficos residuales y los parámetros estadísticos que se obtuvieron del modelo indicaron que si bien el modelo tenía las pruebas
- El modelo obtenido logra representar aproximadamente un 80% de la realidad, por lo que permite a la empresa generar estándares de garantía con su cliente a la hora de cumplir con los tiempos requeridos.
- Al usar el modelo y disminuir el porcentaje de multas en las que se incurre por proyecto, el ROI de la empresa se ve afectado de manera que le permite subir hasta 10 puntos respecto al actual.

8. RECOMENDACIONES

- Los datos recopilados durante el último año contienen mucho ruido, lo cual dificulta mucho la estandarización de un proceso, la información que retroalimenta la empresa de cada proyecto debe ser más específica, para que de esta manera se tengan más ítems con una serie de datos típicos y no un solo ítem que tenga muchos datos atípicos por no tomar en cuenta su diversificación.
- La Linieria es la actividad critica a la hora de hablar de tiempo, por lo que la empresa DR & MR debe certificar que cada uno de los técnicos dedicados a esta actividad que tengan cumplan con el requisito del contrato con el cliente, que es tender 500 mts de fibra diarios, de esta manera en la asignación de personal es más ajustada a la realidad.
- La empresa debe empezar a usar un modelo para la asignación de recursos, si bien no va a ser el arreglo total para todos los proyectos, este modelo le va a permitir empezar a entregar proyectos dentro de los tiempos e ir reduciendo el porcentaje de multas.

9. BIBLIOGRAFIA

- Arroyave, C. (2012). <http://repository.lasallista.edu.co>. Recuperado el 01 de Marzo de 2017, de <http://repository.lasallista.edu.co/dspace/bitstream/10567/714/1/INFORME%20DE%20PRACTICA%20CAROLINA%20GONZALEZ%20ARROYAVE.pdf>
- Casari, A. (2012). *LOS 5 ERRORES FATALES DE LOS ANALISTAS DE PROCESOS*. casari-asoc.com. Recuperado el 18 de Febrero de 2017
- Freivalds, N. . (2004). *Ingenieria Industrial Metodos, estándares y diseño del trabajo*. Mexico : Alfaomega.
- Lopez, B. S. (12 de 2006). Obtenido de *ingenieriaindustrialonline*: <https://www.ingenieriaindustrialonline.com/>
- Mongomery, D., & Runger, G. (2003). *Applied Statistics and Probability for Engineers*. Wiley. Recuperado el 10 de Febrero de 2017
- Moore, D. S. (2004). *Estadística aplicada básica*. Purdue University. Recuperado el 10 de Febrero de 2017
- Nuñez, A. (1992). *Estadística básica para planificador*. Mexico: Siglo Ventiuno. Recuperado el 11 de Febrero de 2017
- Salazar, B. (2016). *IngenieriaIndustrialOnline*. Recuperado el 11 de Febrero de 2017, de <https://www.ingenieriaindustrialonline.com/herramientas-para-el-ingeniero-industrial/pron%C3%B3stico-de-ventas/regresi%C3%B3n-lineal/>

ANEXO A MATRIZ DE DATOS POR PROYECTO

NUMERO DE PROYECTO	INFORMACIÓN TÉCNICA DEL PROYECTO							INFORMACIÓN SOBRE LAS INVERSIONES DE LA EMPRESA PARA REALIZAR EL PROYECTO					RESULTADOS DEL PROYECTO	
	CIUDAD	CLIENTE	CANTIDAD DE CABLE	TIPO DE CABLE	TIPO DE APOYO	CANTIDAD DE ENRUTAMIENTO	PUNTO DE PARTIDA	HORAS DISEÑADOR	HORAS INGENIERO	HORAS LINIERO	HORAS EMPALMADOR	HORAS CONDUCTOR	TIEMPO TOTAL	ROI DEL PROYECTO
1	PEREIRA	ACCEDO SANTANDER	800	48	1	2	1	8	40	16	8	16	80	0,12
2	CAJAMARCA	BANCOLOMBIA	180	6	1	4	0	8	20	16	8	8	40	0,09
3	BUCARAMANGA	AMWAY	300	6	1	2	0	4	20	24	8	0	40	0,05
4	SALAMINA	MEDICINA LEGAL	300	6	1	5	0	8	32	24	16	16	56	0,1
5	HONDA	MEDICINA LEGAL	150	12	1	3	1	4	20	16	6	8	40	0,11
6	PEREIRA	UNILEVER	232	12	1	3	1	4	20	16	16	16	56	0,05
7	GUADUAS	BANCAMIA	1500	48	1	2	0	8	40	32	16	8	64	0,12
8	GIRARDOT	UDEC	1850	12	1	3	1	8	40	48	24	8	56	0,15
9	CUCLUTA	VILLA DEL ROSARIO	380	6	1	5	1	4	20	16	8	0	32	0,07
10	AGUACHICA	POSTOBON	313	12	1	0	0	8	20	16	8	16	40	0,08

30	DOSQUEBRADAS	COLUMBUS NETWORKS STECKERL ACEROS S.A.S	740	12	1	2	1	8	32	16	16	16	56	0,09
31	GINEBRA	LEVEL 3 SOLUCIONES INTEGRALES SAN PEDRO	1670	48	0	4	1	8	32	16	32	16	56	0,1
32	PEREIRA	NETWORKS PROFAMILIA	1180	48	1	3	0	8	32	16	32	16	56	0,1
33	BUCARAMANGA	COMPUDEC OUTSORCING	700	12	1	1	0	8	24	16	8	6	48	0,1
34	SAN ALBERTO	INTERNET DEDICADO PROYECTO WIFI	1130	12	0	4	1	4	32	64	32	8	48	0,13
35	PEREIRA	LEVEL 3 SALUD TOTAL	500	6	1	8	1	4	24	16	40	16	40	0,11
36	CUCLUTA	IFX NETWORKS AGENCIA NACIONAL DE TIERRAS	500	12	0	4	0	4	20	16	24	6	32	0,11
37	SAN ALBERTO	FINCA	200	6	1	2	0	4	20	16	8	8	24	0,09
38	OCAÑA	IFX NETWORKS CNR	250	6	1	2	0	4	24	16	8	16	32	0,1
39	SE SAHAGUN	ELECTRICARIBE	205	6	0	1	0	4	20	24	8	8	24	0,09
100	SE SAMPUES	ELECTRICARIBE	480	6	0	3	0	4	20	16	16	8	24	0,08

Anexo C Tabla de residuales modelos escogidos

<i>Observation</i>	<i>Predicted T. TOTAL</i>	<i>Standard Residuals</i>		<i>Percentile</i>	<i>T. TOTAL</i>
1	33,72770005	1,011821906		0,588235	24
2	34,63143445	0,866035151		1,764706	24
3	46,02191667	1,609623804		2,941176	24
4	33,4000933	1,064670096		4,117647	24
5	56,27029633	1,246924345		5,294118	24
6	56,27029633	-0,043603104		6,470588	24
7	34,63143445	-0,424492298		7,647059	32
8	32,4963589	1,210456851		8,823529	32
9	55,94268957	0,009245085		10	32
10	46,02191667	-2,261958544		11,17647	32
11	73,72824572	-1,569321016		12,35294	32
12	56,34938221	-1,346888366		13,52941	32
13	77,58044093	-0,900214025		14,70588	32
14	56,34938221	-1,346888366		15,88235	32
15	49,13981249	1,106657536		17,05882	32

.
.
.

70	32,16875214	1,26330504		81,76471	56
71	33,70792858	1,015011359		82,94118	56
72	36,6772714	0,536009058		84,11765	56
73	47,90847134	1,305292481		85,29412	56
74	65,26756337	-0,204478166		86,47059	56
75	46,02191667	1,609623804		87,64706	64
76	46,02191667	1,609623804		88,82353	64
77	46,02191667	1,609623804		90	64
78	38,5440546	1,525394638		91,17647	64
79	46,92565106	0,173309599		92,35294	64
80	36,6772714	0,536009058		93,52941	72
81	33,70792858	-0,27551609		94,70588	72
82	33,4000933	-1,516384803		95,88235	72
83	36,6772714	-0,754518391		97,05882	72
84	33,4000933	-1,516384803		98,23529	72
85	33,4000933	-1,516384803		99,41176	80
