attrition-and-per	rom work and tenurion and areas of en	re encompas nhancement i	ss demographics in HR practice.	s of the employees	ngagement in the wo which make the data  /hr-analytics-e	a set useful in
1.2. Explain the  1. EmployeeID: The re 2. FirstName: The firs 3. LastName: The last	features: corded date is Unic	oyee.	er for each emplo	oyee.		
4. <b>Gender:</b> The gende 5. <b>Age:</b> The age of the 6. <b>BusinessTravel:</b> Th 7. <b>Department:</b> The d 8. <b>DistanceFromHom</b> 9. <b>State:</b> The state in v	employee. e frequency of bus epartment in which e(KM): The distand	n the employ ce between t	ee works.		ce in kilometers.	
10. Ethnicity: The ethnicity: The annual: 12. Salary: The annual: 13. StockOptionLevel: 14. OverTime: Whether 15. HireDate: The date	marital status of the salary of the emploon The level of stock of the employee work the employee was	ne employee. oyee. options gran ks overtime hired.	ited to the emplo (Yes/No).	oyee.		
16. Attrition: Whether to 17. YearsAtCompany: 18. YearsInMostRecen 19. YearsSinceLastPro 20. YearsWithCurrMan	The number of year tRole: The number motion: The numb	rs the emplo of years the per of years s	oyee has been w e employee has l since the employ	been in their most i yee's last promotion	n.	
# Importing library import numpy as np import pandas as po import seaborn as s import matplotlib.p from scipy import s from sklearn.cluste from sklearn.prepro	ns yplot <b>as</b> plt tats r <b>import</b> KMeans		aler			
<pre>from sklearn.linear from sklearn.metric from sklearn.model_ from scipy.stats in import statsmodels. from sklearn.model_ from sklearn.preprof from sklearn.pipeli</pre>	selection import zscore api as sm selection import zscore api as sm selection import zscore import z	<pre>quared_err t train_te  t train_te  Polynomial</pre>	or st_split, KFo		core	
<pre>from scipy.interpol  # 1.3. Loading the df = pd.read_csv('E)  # 1.4. Show the load </pre>	dataset		pline			
<pre>df.head()</pre>	me LastName Ge	nder Age E	BusinessTravel I Some Travel	<b>Department Distan</b> Sales	ceFromHome (KM) State	<b>Ethnicity Marit</b> White
<ol> <li>CBCB-9C9D Leon</li> <li>95D7-1CE9 Ahr</li> <li>47A0-559B Ermentre</li> </ol>	ned Sykes	Male 38  Male 43  Non-inary 39	Some Travel  Some Travel	Sales  Human Resources  Technology	23 CA 29 CA 12 IL	White  Asian or Asian  American  White
4 42CC-040A St  5 rows × 23 columns  # Show the loaded in df.info()		emale 29	Some Travel	Human Resources	29 CA	White
<pre><class #="" 'pandas.core="" (total="" 1470="" column<="" columns="" data="" end="" rangeindex:="" td=""><td>ries, 0 to 1469 23 columns): Non-N  1470 1470</td><td></td><td>Dtype object object object</td><td></td><td></td><td></td></class></pre>	ries, 0 to 1469 23 columns): Non-N  1470 1470		Dtype object object object			
3 Gender 4 Age 5 BusinessTravel 6 Department 7 DistanceFromHor 8 State 9 Ethnicity 10 Education 11 EducationField	1470 1470 1470 ne (KM) 1470 1470 1470	non-null non-null non-null non-null non-null non-null non-null non-null non-null	object int64 object object int64 object object int64 object			
Jobrole Jobrol	1470 1470 1470 1470 1470 1470 1470	non-null non-null non-null non-null non-null non-null non-null non-null	object object int64 int64 object object object int64			
20 YearsInMostRece 21 YearsSinceLastI 22 YearsWithCurrMa dtypes: int64(9), of memory usage: 264.3- ]: # 1.5 Show some tree df.describe()	Promotion 1470 anager 1470 oject(14) - KB	non-null non-null non-null	int64 int64 int64			
count 1470.000000 mean 28.989796	1470.000000 16 22.502721	2.912925	1470.000000 112956.497959	1470.000000 0.793878	YearsAtCompany Ye 1470.000000 4.562585	1470.00000 2.29319
std       7.993055         min       18.000000         25%       23.000000         50%       26.000000         75%       34.000000	12.811124 1.000000 12.000000 22.000000 33.000000	1.024165 1.000000 2.000000 3.000000 4.000000	103342.889222 20387.000000 43580.500000 71199.500000 142055.750000	0.852077 0.000000 0.000000 1.000000 1.000000	3.288048 0.000000 2.000000 4.000000 7.000000	2.539093 0.000000 0.000000 1.000000 4.000000
max 51.000000  Section 2  Lasso Regulari	45.000000 <b>zation</b>	5.000000	547204.000000	3.000000	10.000000	10.00000
# Section 2 # 2.1. Prepare data # Filter the list of features2 = ['Age', # Define features a	of for Linear reg of features 'DistanceFromHo				, 'YearsAtCompany	y <b>', '</b> YearsInMostF
<pre>X = df[features2] y = df['Salary']  # Standardize the description scaler = StandardSd X_scaled = scaler.f</pre>	data :aler() :it_transform(X)					
<pre># 2.3. Train linear linear_reg = Linear linear_reg.fit(X_sc  # 2.4. Build a lass alpha = 100 lasso_reg = Lasso(a lasso_reg.fit(X_sca</pre>	Regression() caled, y) co regularization	n model an	d choose a va	lue for alpha hy	yper-parameter.	
<pre>lasso_cv = LassoCV lasso_cv.fit(X_scal alpha = lasso_cv.al print(f"Optimal alp  # Calculate MSE for y pred linear = linear</pre>	ed, y)  pha  ha for Lasso: {  both models	alpha}")	5-fold cross-	validation		
y_pred_lasso = lass  mse_linear = mean_s  mse_lasso = mean_so  print(f"Mean Square  print(f"Mean Square	equared_error(y, guared_error(y, guared_error(	_scaled) y_pred_li y_pred_las ear Regres	so) sion: {mse_li:		sso}")	
<pre># 2.5. Obtain the o print("\nLinear Rec for feature, coef i     print(f"{feature  print("\nLasso Regr for feature, coef i</pre>	ression Coefficant zip (features 2 re): {coef}")	ients (Sca., linear_re	<pre>led):") eg.coef_): ed):")</pre>			
optimal alpha for La Mean Squared Error: Mean Squared Error: Linear Regression Co Age: -1.006835269345	asso: 103.307732 For Linear Regre For Lasso Regres	ession: 8.9 ssion (alph			.00000000224	
DistanceFromHome (KN Salary: 103307.7325! StockOptionLevel: 1 YearsAtCompany: 2.54 YearsInMostRecentRol YearsSinceLastPromoty YearsWithCurrManage: 10300000000000000000000000000000000000	995339 .250555214937776 .6585164964199e- .e: 2.9103830456 .cion: 3.63797880 .c: 1.81898940354	3e-12 11 733704e-11 7091713e-1 58565e-11				
Lasso Regression Coe Age: 0.0 DistanceFromHome (KN Salary: 103207.73259 StockOptionLevel: -0 YearsAtCompany: 0.0 YearsInMostRecentRol YearsSinceLastPromos	1): 0.0 9953388 ).0	.ed):				
YearsWithCurrManage						
2.6. Analyze the control of the coefficients in the line means for this model, say only salary has a non-zeron and completely ignores of the feature selection don't help that much. Late the mean squared error sign that the model is over the coefficient of the coefficients of the coefficients of the coefficients of the coefficients in the line means of the coefficients o	lifferences be near regression more lary is very important to coefficient where other features.  In linear regression only for linear regression only erfitting, meaning in the training data at ing.  In offers highly accurate regression simplification.	ant where other features all the fly picks important might not varies perfectly.	tremely small van her features real ures are zero. Lan features without rtant features by zero, which mea work well on new But because it is aining, but it car	alues for most featurely don't matter must asso regression also any penalty, so it by making the coefficients it fits the training, unseen data. Last gnores unimportant on be overfitting becautives that don't a	ch. On the other hand thinks the same that the same that the same that the same that the same everything ever cient zero.  If data perfectly. How so regression has high the features, it may persease it includes all feadd much value. It misses the sadd much value.	d, for lasso regression salary is important and salary is important in if some features wever, this can be a gher Mean Squared form better on new eatures, even the
2.6. Analyze the control of the coefficients in the lift means for this model, satisfied and completely ignores an	lifferences be near regression modern to coefficient where other features.  In linear regression only for linear regression only erfitting, meaning is the training data at ing.  In offers highly accurate regression simplifies more likely to work.	ant where other eather other features all the five picks important and provided the fit on the fit	tremely small van her features real ures are zero. Lan features without rtant features by zero, which mea work well on new But because it is aining, but it car	alues for most featurely don't matter must asso regression also any penalty, so it by making the coefficients it fits the training, unseen data. Last gnores unimportant on be overfitting becautives that don't a	ch. On the other hand thinks the same that the same that the same that the same that the same everything ever cient zero.  If data perfectly. How so regression has high the features, it may persease it includes all feadd much value. It misses the sadd much value.	d, for lasso regression salary is important and salary is important in if some features wever, this can be a gher Mean Squared form better on new eatures, even the
2.6. Analyze the control of the coefficients in the lift means for this model, satisfied and completely ignores are for the feature selection don't help that much. Lat the mean squared error sign that the model is over the feature selection and the model is over the means it doesn't find the data by avoiding overfitt.  Overall, linear regression unimportant ones. Lasso data as perfectly, but it's section 4  K-fold Cross  # Section 4 - K-foid # 4.1. Set up your features2 = ['Age', X = df[features2] y = df['Salary'] # # 4.2. Initialize to	lifferences be near regression modern lary is very important to coefficient where other features.  In linear regression only for linear regression only for linear regression erfitting, meaning in the training data as ing.  In offers highly accumulated accumulated regression simplifies more likely to work to more likely to work to regression simplifies and cross-validated data for K-fold bistanceFromHead data for K-fold to histanceFromHead data for K-fold data for K-fold to histanceFromHead data for K-fold data fo	odel show extent where other features all the fly picks imported it might not very perfectly.  I cross-value (KM) ',	tremely small values features are zero. La features without rtant features by zero, which meawork well on new But because it is aining, but it carel by removing for new data because it is new data because i	alues for most featurely don't matter mulasso regression also any penalty, so it is any penalty, so it is y making the coefficiens it fits the training, unseen data. Last gnores unimportant to be overfitting because it avoids overfitting use it avoids overfit	ch. On the other hand thinks the same that seeps everything ever cient zero.  If data perfectly. How so regression has high the features, it may perfectly ause it includes all feadd much value. It mitting.	d, for lasso regression at salary is important and salary is important and if some features wever, this can be an agher Mean Squared form better on new eatures, even the light not fit the training
2.6. Analyze the control of the coefficients in the lift means for this model, satisfied and completely ignores and completely ignores and completely ignores and completely ignores are for the feature selection don't help that much. Lat the mean squared error sign that the model is over the feature selection and the model is over the means it doesn't fir data by avoiding overfitt.  Overall, linear regression unimportant ones. Lasso data as perfectly, but it's section 4  K-fold Cross  # Section 4 - K-form # 4.1. Set up your features2 = ['Age', X = df[features2] y = df['Salary'] #	lifferences be near regression modern lary is very important to coefficient where other features.  In linear regression only for linear regression erfitting, meaning in the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.	ion  ion  ion  ion  ion  ion  ion  ion	tremely small values features are zero. La features without rtant features by zero, which meawork well on new But because it is aining, but it carel by removing for new data because it is new data because it is attention.	alues for most feature lly don't matter must asso regression also any penalty, so it by making the coefficients it fits the training, unseen data. Last gnores unimportant to be overfitting bed eatures that don't a use it avoids overfit.  'StockOptionLever's associated as a second control of the coefficients are in the coefficients as a second coefficient of the coefficients are in the coefficients as a second coefficient of the coefficients are in the coefficients as a second coefficient of the coefficients are coefficients.	ch. On the other hand thinks the same that t	d, for lasso regression at salary is important on if some features evever, this can be a gher Mean Squared form better on new eatures, even the light not fit the training the same of the
2.6. Analyze the control of the coefficients in the lift means for this model, satisfied only salary has a non-zer and completely ignores of the feature selection don't help that much. Late the mean squared error sign that the model is over the feature selection don't help that much. Late the means it doesn't fir data by avoiding overfitt.  Overall, linear regression unimportant ones. Lasso data as perfectly, but it's section 4  K-fold Cross  Section 4  K-fold Cross  # Section 4 - K-food # 4.1. Set up your features2 = ['Age',  X = df[features2] y = df['Salary'] #  # 4.2. Initialize the linear reg = Linear reg	lifferences be near regression modern lary is very important to coefficient where other features.  In linear regression only for linear regression only for linear regression erfitting, meaning is the training data as ing.  In offers highly accumulate regression simplifies more likely to work the modern larger of the linear regression of the lin	cross-val ome (KM)',  one (KM)	tremely small value her features real ures are zero. La features without rtant features by zero, which mea work well on new But because it is aining, but it carel by removing for new data because it is new data because it is at a test of the control of the cont	alues for most feature lly don't matter must asso regression also any penalty, so it by making the coefficients it fits the training, unseen data. Last gnores unimportant in the overfitting bed eatures that don't associate it avoids overfit use it avoids overfit when the coefficients with the coefficients w	ch. On the other hand thinks the same that t	d, for lasso regression at salary is important on if some features evever, this can be a gher Mean Squared form better on new eatures, even the light not fit the training the same of the
2.6. Analyze the control of the coefficients in the liming means for this model, say only salary has a non-zer and completely ignores of the feature selection don't help that much. Late the model is one and the complete of the complete of the feature selection don't help that much. Late the model is one and the complete of the compl	lifferences be near regression modern lary is very important to coefficient where other features.  In linear regression only for linear regression only for linear regression erfitting, meaning in the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the Linear Regression ()  In offers highly accumulate the Linear	ion  ion  ion  ion  ion  ion  ion  ion	tremely small value her features real ures are zero. La features without rtant features by zero, which mea work well on new But because it is aining, but it carel by removing for new data because it is new data because it is at a test of the control of the cont	alues for most feature lly don't matter must asso regression also any penalty, so it by making the coefficients it fits the training, unseen data. Last gnores unimportant in the overfitting bed eatures that don't associate it avoids overfit use it avoids overfit when the coefficients with the coefficients w	ch. On the other hand thinks the same that t	d, for lasso regression at salary is important on if some features evever, this can be a gher Mean Squared form better on new eatures, even the light not fit the training the same of the
2.6. Analyze the control of the coefficients in the lift means for this model, satisfies only salary has a non-zer and completely ignores of the feature selection don't help that much. Late the mean squared error sign that the model is one to the feature selection don't help that much. Late the means it doesn't fir data by avoiding overfitt.  Overall, linear regression unimportant ones. Lasso data as perfectly, but it's section 4  K-fold Cross  Section 4  K-fold Cross  # Section 4 - K-foid # 4.1. Set up your features2 = ['Age', X = df[features2] y = df['Salary'] #  # 4.2. Initialize the linear reg = Linear # 4.3. Define a 5-in kf = KFold(n_splits) # 4.5. Manually look fold = 1 mse_scores = [] # for train_index, test y_train, y_test # Train the model linear_reg.fit # Predict on the y_pred = linear # Calculate MSI mse = mean_squamse_scores.apped print(f"Fold { fold += 1 } # Print results average_mse = np.me variance_mse =	lifferences be near regression molary is very importation coefficient where the coefficient where other features.  Ilinear regression only for linear regression only for linear regression erfitting, meaning is the training data as ing.  In offers highly accurate regression simplifies more likely to work the Linear Regression ()  In offers highly accurate regressi	codel show extent where other feature and where other feature was all the fly picks important and picks important fit on the case perfectly.  In the company of the company	tremely small value her features real ures are zero. La features without rtant features by zero, which measure work well on new But because it is aining, but it carel by removing for new data because it is new data because it is attacted to the control of the c	alues for most feature asso regression also regression also read any penalty, so it by making the coefficients it fits the training, unseen data. Last gnores unimportant to be overfitting bed eatures that don't a use it avoids overfit where the death of the death o	ch. On the other hand thinks the same that the thinks the same that the thinks the same that the telephone everything eve	d, for lasso regression at salary is important on if some features evever, this can be a gher Mean Squared form better on new eatures, even the light not fit the training the same of the
2.6. Analyze the control of the coefficients in the liming means for this model, say only salary has a non-zero and completely ignores of the feature selection don't help that much. Late the mean squared error sign that the model is over the feature selection don't help that much. Late the model is over the feature sit doesn't find the data by avoiding overfitt.  Overall, linear regression unimportant ones. Lasso data as perfectly, but it's section 4  K-fold Cross  **Section 4**  K-fold Cross  **A **Loset up your features 2 = ['Age', X = df[features 2] y = df['Salary'] #  **# 4.2. Initialize the linear reg = Linear #  **# 4.3. Define a 5-1 kf = KFold (n_splits #  **# 4.5. Manually lost fold = 1 mse_scores = [] #  for train_index, test y_train, y_test #  **# Train the model inear_reg.fit the second inear_reg.fit #  **# Predict on the y_pred = linear #  **# Calculate MSE mse = mean_squamse_scores.apped print (f"Fold (fold += 1) #  **# Print results average_mse = np.me variance_mse =	lifferences be near regression moder lary is very important of coefficient where other features.  In linear regression only for linear regression only for linear regression erfitting, meaning in the training data asing.  In offers highly accurate the training data asing.  In offers highly accurate the linear Regression simplifies more likely to work the Linear Regression ()  In the Linear Regression of the Linear Regression ()  In the Linear Regression of the Linear Regression ()  In the Linear Regression of the Linear Regression ()  In the Linear Regression of the Linear Regression ()  In the Linear Regression of	codel show extent where other feature and where other feature with the second state of	tremely small value features real are zero. Last features without real features by zero, which means work well on new But because it is aining, but it cared by removing for new data because it is at a feature features.  idation 'Education',  idation 'Education',  testate=42)  features without real features by zero, which means work well on new But because it is aining, but it cared by removing features are also and the cared by removing features.  idation 'Education',  testate=42)  features without real features by zero, which means work well on new But because it is aining, but it cared by removing features.  idation 'Education',  testate=42)  features without real features by zero, which means work well on new But because it is aining, but it cared by removing features.  idation 'Education',  testate=42)  features without real features by zero, which means work well on new But because it is aining, but it cared by removing features.  idation 'Education',  testate=42)  features without real features by zero, which means work well on new But because it is aining, but it cared by removing features.  idation 'Education',  and the state are also and the sta	alues for most feature for most feature for matter mutures asso regression also any penalty, so it is any penalty, so it is any making the coefficient it is the training, unseen data. Last gnores unimportang the overfitting bed eatures that don't a use it avoids overfit where the dexistance of the matter fold dexistance of the matter	ch. On the other hands thinks the same that	d, for lasso regression at salary is important on if some features evever, this can be a gher Mean Squared form better on new eatures, even the light not fit the training the same of the
2.6. Analyze the C  The coefficients in the lime means for this model, satisfied only salary has a non-zer and completely ignores and completely ignores and completely ignores are for the feature selection don't help that much. Late the model is over the feature selection and the property of the feature selection don't help that much. Late the model is over the feature selectly, but it's the section 4  K-fold Cross  Section 4  K-fold Cross  K-fold Cross  Section 4  K-fold Cross  Section 4  K-fold Cross  A Section 4 - K-fold  # 4.1. Set up your features 2 = ['Age', X = df[features 2] Y = df['Salary'] #  # 4.2. Initialize the linear reg = Linear features 2 = L	lifferences be near regression moderly is very importation coefficient where the features.  Ilinear regression only for linear regression erfitting, meaning it the training data at ing.  In offers highly accumore likely to work the training data at ing.  In offers highly accumore likely to work the training data at the Linear Regression ()  In the Linear Regress	codel show extent where other feature and where other feature are other feature uses all the five picks important as perfectly.  In the fit on the fiest the model of the fit on the fiest the model of the fiest the model of the fiest form of the f	tremely small value features real ures are zero. Last features without real transfeatures by zero, which means work well on new But because it is aining, but it cared by removing for new data because it is at the state of the contract of the state of the s	alues for most feature for most feature for matter mutures asso regression also any penalty, so it is any penalty, so it is any making the coefficient it is the training, unseen data. Last gnores unimportang the overfitting bed eatures that don't a use it avoids overfit where the dexistance of the matter fold dexistance of the matter	ch. On the other hands thinks the same that	d, for lasso regression at salary is important on if some features evever, this can be a gher Mean Squared form better on new eatures, even the light not fit the training the same of the
2.6. Analyze the C  The coefficients in the limeans for this model, so only salary has a non-zer and completely ignores of the feature selection don't help that much. La  The mean squared error sign that the model is over the feature selection don't help that much. La  The mean squared error sign that the model is over the feature selection don't help that much. La  The mean squared error sign that the model is over the feature selection don't help that much. La  The mean squared error sign that the model is over the feature selection don't help that much. La  The mean squared error sign that the model is over the feature selection don't help that model is over the feature selection don't help that the feature selection don't help that feature selection data as perfectly, but it's  Section 4  K-fold Cross  Section 4  K-fold Cross  # Section 4 - K-fold  # 4.1. Set up your features2   Year featur	lifferences be near regression modern is very importation coefficient where other features.  Ilinear regression only for linear regression only for linear regression erfitting, meaning in the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training data as ing.  In offers highly accumulate the training and the tr	codel show extent where other features all the five picks imported in the picks imported	tremely small value features real are zero. Lad features without real features by zero, which means work well on new But because it is aining, but it care by removing for new data because it is aining, but it care by removing for new data because it is at a state for the control of the cont	alues for most feature for most for most feature f	ch. On the other hands thinks the same that thinks the same that the same that the seeps everything	d, for lasso regression at salary is important on if some features evever, this can be a gher Mean Squared form better on new eatures, even the light not fit the training the same of the
2.6. Analyze the C  The coefficients in the limeans for this model, so only salary has a non-zero and completely ignores of the feature selection don't help that much. Late the model is over the feature selection don't help that much. Late the model is over the feature selection don't help that much. Late the model is over the feature selection don't help that much. Late the model is over the feature selection don't help that much. Late the model is over the feature selection don't help that much. Late the model is over the feature selection don't help that the model in the feature selection don't help that the model in the feature selection don't help that the feature	lifferences be near regression moder is very importation coefficient where other features.  Ilinear regression only for linear regression only for linear regression only for linear regression erfitting, meaning it the training data at ing.  In offers highly accurate regression simplifies more likely to work the second of t	codel show extent where other features and where other features are other features are other features and the form of the price of the fit on the price of the fit on the fit of	tremely small value features real are are zero. Last features without real features by zero, which means work well on new But because it is aining, but it care by removing for new data because it is at a state and features by removing features by removing features by removing featu	alues for most featured lly don't matter must asso regression also any penalty, so it was it fits the training was any penalty, so it was any penalty and the training was any pena	ch. On the other hands the same that the sam	d, for lasso regression at salary is important and sal
2.6. Analyze the C  The coefficients in the limeans for this model, sa only salary has a non-ze and completely ignores and completely ign	lifferences be mear regression mo lary is very importate ro coefficient where other features.  Ilinear regression only for linear regression sso regression only for linear regression erfitting, meaning it the training data at ing.  In offers highly accurate regression simplifies more likely to world  In offers highly accurate regression simplifies more likely to world  In offers highly accurate regression simplifies more likely to world  In offers highly accurate regression simplifies regression simplifies regression of the linear regres regression of the linear regre	codel show extent where other features all the five other features all the five picks important and the process of the first on the first the model of the first the model of the first the model of the first the fold of the first the first the fold of the first the first the fold of the first the	tremely small value features real are zero. Last features without real features without real features by zero, which meal work well on new But because it is aining, but it carel by removing for new data because it is aining, but it carel by removing for new data because it is aining, but it carel by removing for new data because it is aining. The search fold ets for the c.iloc[test_indic] iloc[test_indic] iloc[	alues for most featured lly don't matter must asso regression also any penalty, so it was it fits the training was any penalty, so it was any penalty and the training was any pena	ch. On the other hands the same that the sam	d, for lasso regression at salary is important and sal
2.6. Analyze the C  The coefficients in the lime means for this model, sa only salary has a non-ze and completely ignores of the feature selection don't help that much. Late the model is over the feature selection don't help that much. Late the model is over the feature selection don't help that much. Late the model is over the feature selection don't help that much. Late the model is over the feature selection don't help that model is over the feature selection don't help that model is over the feature selection don't help that the model in the feature selection don't help that help the feature selection don't help that help th	lifferences be near regression mo lary is very importer to coefficient where other features.  Ilinear regression only for linear regression erfitting, meaning if the training data at ing.  In offers highly accurate the training data at ing.  In offers highly accurate the training data at ing.  In offers highly accurate the training data at ing.  In offers highly accurate the training data at ing.  In offers highly accurate the training and at a for K-fold Interpolation  In the Linear Regresion ()  In	codel show extent where other features all the five other features all the five picks important and the process of the first the model of the first the form of the first the first the model of the first the first the model of the first the firs	tremely small value features real are zero. Last features without real features without real features by zero, which means work well on new But because it is aining, but it carel by removing for new data because it is at the state of the control of the state of the control of the state of the control of the state of	alues for most feature ly don't matter must asso regression also asso regression also any penalty, so it is any penalty, so it is y making the coefficients if its the training, unseen data. Last gnores unimportant in the overfitting becautes that don't as use it avoids overfit avoids overfit was it avoids overfit avoids overfit avoids overfit was accessed by the season of the s	ch. On the other ham of thinks the same that thinks the same that the thinks the same that the thinks the same that the teleps everything every	d, for lasso regression at salary is important and salary is important in if some features evever, this can be a gher Mean Squared form better on new eatures, even the ight not fit the training early, 'YearsInMore and the salary', 'YearsInMore and the salary', 'YearsInMore and the salary', 'YearsInMore and the salary is important and th
2.6. Analyze the C  The coefficients in the lime means for this model, so only salary has a non-zero and completely ignores. For the feature selection don't help that much. Late the mean squared error sign that the model is over the data by avoiding overfitt. Overall, linear regression unimportant ones. Lassed data as perfectly, but it's section 4  K-fold Cross  K-fold Cross  K-fold Cross  K-fold Cross  A K-fold Cross  Section 4  K-fold Cross  K-fold Cross  A La C Initialize  Fold 1 Section  A A 1 A 2 Refered  A 1 A 2 Refered  A 2 Initialize  A 2 Initialize  A 3 Define a 5-1  A 5 Manually loc  for train_index  A 2 Initialize  A 4 2 Initialize  A 3 Define a 5-1  A 5 Manually loc  for train_index  A 2 Initialize  A 5 Initialize  A 4 2 Initialize  A 6 Initialize  A 6 Initial	lifferences be near regression molary is very importate to coefficient where other features.  In linear regression only service the training data as ing.  In offers highly accurate the training data as ing.  In offers highly accurate the training data as ing.  In offers highly accurate the training data as ing.  In offers highly accurate the training data as ing.  In offers highly accurate the training and as ing.  In offers highl	del show extent where other feature and where other feature are ot	tremely small value features real ures are zero. Last are zero, which means work well on new But because it is aining, but it care let by removing for new data because it is aining, but it care let by removing for new data because it is at a serior in the control of the contr	alues for most feature and the second street of the	ch. On the other ham be thinks the same that thinks the same that the thinks the same that the thinks the same that the teeps everything every cient zero.  If data perfectly. How the same it includes all features, it may perfectly the thinks the thinks the same it includes all feated much value. It misting.  If the interpret is a same that the teeps are the same and the same are	d, for lasso regression at salary is important and salary is important in if some features wever, this can be a gher Mean Squared form better on new eatures, even the ight not fit the training the dependence of the salary of the december of the salary of
2.6. Analyze the C  The coefficients in the limeans for this model, so only salary has a non-zee and completely ignores. For the feature selection don't help that much. La The mean squared error sign that the model is over the feature selection don't help that much. La The mean squared error sign that the model is over the feature selection of the property of the feature selection of the featu	description of the current of the cu	del show eximit where of the other feature is all the five picks important in the sperfectly.  In the fit of the sperfectly is a sperfectly.  In the fit of the sperfectly is a sperfectly in the specific specifi	tremely small value features real ures are zero. La features without rtant features by zero, which mea work well on new But because it is aining, but it care by removing for new data because it is aining, but it care by removing for new data because it is aining that it stands are also for the c. iloc[test_insiloc	nt with the five folce hosen are close to learn make it easy trees of the series of th	ch. On the other hands the other hands the same that the s	the highest error rate erage MSE is in the gentures, even the ight not fit the training the folds. Overage may', 'YearsInMore equation 4, the delevation of the folds on the folds of the folds of the folds. Overage may the folds. Overage may the folds. Overage may the folds. Overage may the folds.
2.6. Analyze the C  The coefficients in the list means for this model, so only salary has a non-zee and completely ignores. For the feature selection don't help that much. La The mean squared error sign that the model is over the feature selection don't help that much. La The mean squared error sign that the model is over the feature selectly, but it's data by avoiding overfith the data by avoiding overfith data by avoiding overfith data by avoiding overfith the features of	differences be near regression mo lary is very importate to coefficient when other features.  Ininear regression only for linear regression erfitting, meaning it the training data at ing.  Inoffers highly accurate regression simplifits the training data at ing.  Inoffers highly accurate regression simplifits the more likely to work  In the Linear Regret Regression ()  In the Linear Regre	del show eximate where other feature and where other feature uses all the five picks important and its might not a sperfectly.  In the fit on the set of the sperfectly are fit on the set of the mode of the sperfectly are folds:  In the fit of the sperfectly are form of the sperfectly are form of the sperfectly are form of the sperfectly are folds:  In the sperfectly are form of the sperfectly are folds:  In the sperfectly are form of the sperfectly.  In the specific transfer of the specific	tremely small value features real ares are zero. Later are zero, which means work well on new But because it is aining, but it carel by removing for new data because it is at a state and it is a state are zero. When the content is a state are zero. It is a state are zero. It is a state are zero. Later	nt with the five folce hosen are close to learn make it easy trees of the series of th	ch. On the other hands the other hands the same that the s	the highest error rate erage MSE is in the gentures, even the ight not fit the training the folds. Overage may', 'YearsInMore equation 4, the delevation of the folds on the folds of the folds of the folds. Overage may the folds. Overage may the folds. Overage may the folds. Overage may the folds.
2.6. Analyze the C  The coefficients in the list means for this model, so only salary has a non-zer and completely ignores. For the feature selection don't help that much. La The mean squared error sign that the model is over the feature selection don't help that much. La The mean squared error sign that the model is over the feature selection of the property of the feature selection of the fe	lifferences be near regression molary is very importation coefficient when other features.  Inlinear regression only for linear regression only for linear regression only for linear regression erfitting, meaning it in the training data a sing.  In offers highly accurate the features of the cornel likely to work the feature of the feat	del show eximite the model's comparation of the manual of the model's comparation of the model's compa	tremely small value features real are sare zero. Later are sare zero. Later are sare zero. Later are sare zero. Later are such as a sare zero, which means work well on new But because it is aining, but it care by removing for new data because it is aining, but it care by removing for new data because it is a sare and the same are sare at the libraries of a sare are are sare at the libraries of a sare are at the libraries of a sare are sare at the libraries of a sare are at the libraries are at the libraries are at the libraries at the libraries are at the libr	alues for most feature alues for most feature to a sort early so it is any penalty, so it is any penalty, so it is a sort early first the training and the coefficients it fits the training and the coefficients are coefficients and the coefficients are coefficients.  If MAKE MODIFICATION were to dearly dexided and and the coefficients are close to dity in the choice of is in for the actual more and the coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficients are coefficients are coefficients are coefficients are coefficients. The coefficients are coefficient	ch. On the other hands the other hands the same that the s	the highest error rate erage MSE is in the gentures, even the ight not fit the training the folds. Overage may', 'YearsInMore equation 4, the delevation of the folds on the folds of the folds of the folds. Overage may the folds. Overage may the folds. Overage may the folds. Overage may the folds.
2.6. Analyze the C  The coefficients in the limeans for this model, so only salary has a non-ze and completely ignores. For the feature selection don't help that much. La  The mean squared error sign that the model is overall, linear regression unimportant ones. Lassed data as perfectly, but it's section 4  K-fold Cross  We fold Cross  The section 4 - K-fold Cross  Section 4 - K-fold Cross  Section 5 - Linear regression unimportant ones. Lassed data as perfectly, but it's section 4 - K-fold Cross  The section 4 - K-fold regression unimportant ones. Lassed data as perfectly, but it's section 4 - K-fold Cross  The section 4 - K-fold regression unimportant ones. Lassed data as perfectly, but it's section 4 - K-fold features 2   Yage', X = df[features2] y = df['Salary'] fold features 2   Yage', X = df[features2] y = df['Salary'] fold features 2   Yage', X = df[features2] y = df['Salary'] fold features 2   Fold features 2	deferences be  mear regression molary is very importar to coefficient when to the features.  In linear regression only for linear regression only for linear regression erifiting, meaning if the training data a ing.  To offers highly accu- tregression simplifit the more likely to work  To a lidatio  To a construction  To a construction	del show eximit where other that where other features all the five the other features are price to the price	tremely small value features real are sare zero. Last features without real features without real features by zero, which mean work well on new But because it is aining, but it care by removing for new data because it is aining, but it care by removing for new data because it is aining that it stands of the control of t	nlues for most feature in the control of the contro	ch. On the other hands the other hands the same that the s	the highest error rate erage MSE is in the gentures, even the ight not fit the training the folds. Overage may', 'YearsInMore equation 4, the delevation of the folds on the folds of the folds of the folds. Overage may the folds. Overage may the folds. Overage may the folds. Overage may the folds.
2.6. Analyze the C The coefficients in the limeans for this model, so only salary has a non-ze and completely ignores: For the feature selection don't help that much. La the selection don't help that much. La the model is one and the selection don't help that much. La the model is one and the selection don't help that much. La the model is one and the selection don't help that much. La the model is one and the selection don't help that much. La the model is one and the selection don't help that much. La the model is one and the selection don't help that much as a perfectly, but it's the selection don't help that the model in the selection don't help that the model in the selection don't help that the selection don't	differences be  mear regression molary is very importative very importative conficient when the features.  Jinear regression only for linear regression only for linear regression or finear regression erfitting, meaning in the training data as ing.  Joffers highly accurate regression simplification of the training data for simplification of the linear regression of the Community	del show eximit where other teath where other feature is all the first picks imported in the picks imported it might not we as perfectly.  Intate fit on the ites the mode is the test in the mode is the pick in	tremely small value features real are sare zero. Last features without real features without real features by zero, which mean work well on new But because it is aining, but it care by removing for new data because it is aining, but it care by removing for new data because it is aining that it stands of the control of t	nlues for most feature in the control of the contro	ch. On the other hands the other hands the same that the s	the highest error rate erage MSE is in the gentures, even the ight not fit the training the folds. Overage may', 'YearsInMore equation 4, the delevation of the folds on the folds of the folds of the folds. Overage may the folds. Overage may the folds. Overage may the folds. Overage may the folds.
2.6. Analyze the C The coefficients in the limeans for this model, so only salary has a non-ze and completely ignores: For the feature selection don't help that much. Late the property of the feature selection don't help that much. Late the property of the feature selection don't help that much. Late the property of the feature selection don't help that much. Late the property of the feature selection don't help that much. Late the property of the feature selection don't help that much. Late the property of the feature selection don't help that the model is one that the model is one that the property of the feature selection don't help that the model is one that the property of the feature selection don't help that the model is one that the feature selection don't help that the feature selection don't help that the model is one that the feature selection don't help that the feature selection d	differences be near regression molary is very importa to coefficient when to coefficient when to coefficient when to ther features.  Ininear regression to six regression only for linear regression teritting, meaning it the training data a ting.  In offers highly accu- tregression simplification to more likely to work the training data a ting.  In offers highly accu- tregression simplification to more likely to work the Linear Regree Regression ()  In offers highly accu- the	del show extent where of the where of the other feature was all the five picks important and its perfectly.  In the first of the set of the set of the mode of the set of the se	tremely small value features real values are zero. Last seatures without real features by zero, which means work well on new But because it is aining, but it carel by removing for new data because it is aining, but it carel by removing for new data because it is aining, but it carel by removing for new data because it is at a section of the control of the con	any penalty, so it is any penalty pe	ch. On the other hands the other hands the same that the s	the highest error rate erage MSE is in the gentures, even the ight not fit the training the folds. Overage may', 'YearsInMore equation 4, the delevation of the folds on the folds of the folds of the folds. Overage may the folds. Overage may the folds. Overage may the folds. Overage may the folds.
2.6. Analyze the C The coefficients in the limeans for this models, sonly salary has a non-ze and completely ignores. The feature selection don't help that much. Last of the property of the feature selection don't help that much. Last of the property of the feature selection don't help that much. Last of the property of the feature selection of	deferences be mear regression molary is very importation of the content when the content end of the content end	del show extent where of the three other feature with the sell of the picks important to the sell of t	tremely small value tremely small value from the features real ares are zero. Ladicates without real features without real features without real features by zero, which meal work well on new But because it is an ing, but it carel by removing for new data because it is an ing, but it carel by removing for new data because it is an ing, but it carel by removing for new data because it is an ing that it stands well on the form the control of the form of the for	alues for most feature and all street and all stree	ch. On the other hands the other hands the same that the s	the highest error rate erage MSE is in the gentures, even the ight not fit the training the folds. Overage may', 'YearsInMore equation 4, the delevation of the folds on the folds of the folds of the folds. Overage may the folds. Overage may the folds. Overage may the folds. Overage may the folds.
2.6. Analyze the C The coefficients in the in means for this model, a only salary has a non-ze and completely ignores: and completely ignores: for the feature selection don't help that much. La The mean squared error signt hat the model is overall, linear regressior unimportant ones. Lasso data as perfectly, but it's Section 4  K-fold Cross  Section 4  K-fold Cross  Section 4  K-fold Cross  A	differences be mear regression mo lary is very importer to coefficient when the reatures.  Ininear regression only seso regression only for linear regression erfitting, meaning in the training data a ing.  Offers highly accu- regression simplification for linear regression erfitting, meaning in the training data a ing.  Offers highly accu- regression simplification for linear regression in defers highly accu- regression simplification for linear regression of the linear regression  Accupation  Offers highly accu- regression simplification for linear regression of the Linear Regression of t	del show eximat where other feature where other feature with the other feature with the picks important the test of the model of the mo	tremely small value treatures readures are zero. Late and treatures by zero, which means work well on new But because it is alining, but it care all by removing for new data because it is alining, but it care all by removing for new data because it is alining, but it care all by removing for new data because it is alining, but it care all by removing for new data because it is alining, but it care all by removing for new data because it is aline to the libraries of its aline that it stands are to all care and its aline that it stands are to all care are are all care are all care are are all care are are all care are are are all care are are are are are are are are are	alues for most featually don't matter musor regression also any penalty, so it is any penalty, so it is y making the coeffins it fits the training, unseen data. Las gnores unimportang the beoverfitting because it avoids overfit ause it avoids overfit ause it avoids overfit ause it avoids overfit ause it avoids overfit aurrent fold dex]  # MAKE MODIFICATE  # MAKE MODIFICAT	ch. On the other hands the other hands the same that the s	the highest error rate erage MSE is in the gentures, even the ight not fit the training the folds. Overage may', 'YearsInMore equation 4, the delevation of the folds on the folds of the folds of the folds. Overage may the folds. Overage may the folds. Overage may the folds. Overage may the folds.
2.6. Analyze the C The coefficients in the limeans for this model, is only salary has a non-ze and completely ignores. For the feature selection don't help that much. La The mean squared error sign that the model is over the feature selection don't help that much. La The mean squared error sign that the model is over the feature sit doesn't fidata by avoiding overfitt data by avoiding overfitt data by avoiding overfitt data by avoiding overfitt data by avoiding overfitt overall, linear regression unimportant ones. Lasso data as perfectly, but it's  Section 4  K-fold Cross  **Cross**  **Section 4 - K-fold**  **Section 4 - K-fold**  **Last up your features2   Y = df   features2   Fold at it X, train, X_test y_train, Y_test y	differences be mear regression mo lary is very importance of coefficient when the features.  In inear regression only for linear regression only for linear regression only for linear regression erfitting, meaning is the training data as ing.  In offers highly accurate the features of the training data as ing.  In offers highly accurate the features of the features	del show extent where other feature where other feature with the other feature with the search form is all the first control of the search fold:  """"" """" """" """"" """"""""""""""	tremely small value freatures without read and services are zero. Late features without read features without read features without read features by zero, which mea work well on new But because it is aining, but it care by removing for an ew data because it is aining, but it care by removing for an ew data because it is aining, but it care by removing for an ew data because it is aining, but it care by removing for an ew data because it is aining, but it care by removing for an ew data because it is aining, but it care by removing for an ew data because it is aining that it stands (rely better performance masses)")  ximum MSE (The search of a formal ized in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performan	allues for most featually don't matter musor regression also any penalty, so it is any penalty and the coefficients of the coefficients and the coefficients are also any penalty and the coefficients are also any penalty and the coefficients are also any penalty an	ch. On the other han of thinks the same than the ceeps everything e	the highest error rate erage MSE is in the gentures, even the ight not fit the training the folds. Overage may', 'YearsInMore equation 4, the delevation of the folds on the folds of the folds of the folds. Overage may the folds. Overage may the folds. Overage may the folds. Overage may the folds.
2.6. Analyze the C  The coefficients in the in means for this model, so only salary has a non-ze and completely ignores:  For the feature selection don't help that much. La The mean squared error signt hat the model is over the selection of the property of the selection of the	lifferences be near regression molary is very importar to coefficient when the coefficient when the features.  In linear regression only for linear regression erfitting, meaning it the training data a ing.  In offers highly accurate the composition of the training data ing.  In offers highly accurate the composition of the composition	del show extent where other to the rest of	tremely small value freatures without read and services are zero. Late features without read features without read features without read features by zero, which mea work well on new But because it is aining, but it care by removing for an ew data because it is aining, but it care by removing for an ew data because it is aining, but it care by removing for an ew data because it is aining, but it care by removing for an ew data because it is aining, but it care by removing for an ew data because it is aining, but it care by removing for an ew data because it is aining that it stands (rely better performance masses)")  ximum MSE (The search of a formal ized in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performance in the libraries claiming that it stands (rely better performan	allues for most feature and the state of the	ch. On the other han thinks the same than the cient zero.  g data perfectly. How so regression has high the features, it may perfectly assessive the same thing.  It is a substituted than the same than the	the highest error rate erage MSE is in the gentures, even the ight not fit the training the folds. Overage may', 'YearsInMore equation 4, the delevation of the folds on the folds of the folds of the folds. Overage may the folds. Overage may the folds. Overage may the folds. Overage may the folds.
2.6. Analyze the C  The coefficients in their means for this model, as only salary has a non-ze and completely ignores.  For the feature selection don't help that much. La  The mean squared error sign that the model is one Error means it doesn't fit data by avoiding overfith Overall, linear regression unimportant ones. Lasse data as perfectly, but it's  Section 4  K-fold Cross  Section 4  K-fold Cross  Section 4  K-fold Cross  Section 4  K-fold Cross  **Section 4 - K-fold  **A-fold Cross  **Section 4 - K-fold  **A-fold Cross  Section 5  **Section 4 - K-fold  **A-fold Cross  **Section 4 - K-fold  **A-fold Cross  Section 5  **Section 4 - K-fold  **A-fold Cross  Section 5  **Section 4 - K-fold  **A-fold Cross  Section 5  **Section 4 - K-fold  **A-fold Cross  Section 6  **A-fold Cross  Section 7  **Section 4 - K-fold  **A-fold Cross  Section 4  K-fold Cross  Section 4  K-fold Cross  Section 4  K-fold Cross  Section 4  K-fold Cross  **Section 4 - K-fold  **A-fold Cross  **A-fold Cross  **Section 4 - K-fold  **A-fold Cross  **A-fold Cross  **A-fold Cross  **Section 4 - K-fold  **A-fold Cross  **A-fold Cross  **A-fold Cross  **A-fold Cross  **A-fold Cross  **Section 4 - K-fold  **A-fold Cross  **A-fold	lifferences be near regression molary is very importar to coefficient when the coefficient when the features.  In linear regression only for linear regression erfitting, meaning it the training data a ing.  In offers highly accurate the composition of the training data ing.  In offers highly accurate the composition of the composition	del show extent where other to the rest of	tremely small value from the statures without restance are zero. Lading and the features without restant features by zero, which mea work well on new But because it is aining, but it carel by removing for new data because it is aining, but it carel by removing for new data because it is aining, but it carel by removing for new data because it is aining the state at a control of t	allues for most feature and the state of the	ch. On the other han of thinks the same than thinks the same than thinks the same than thinks the same than the ceps everything every country of the country	d, for lasso regression at salary is important an if some features wever, this can be a gher Mean Squared form better on new seatures, even the ight not fit the training the highest error rate erage MSE is in the delivering the folds. Over a seature of the folds of the folds of the folds of the folds of the folds. Over a seature of the folds of the folds.
2.6. Analyze the Coefficients in the immension of the imm	inferences be mear regression mo lary is very imports ro coefficient when other features.  In linear regression only for linear regression eritting, meaning it the training data a ing.  Offers highly accu regression simplifi tenore likely to work  In more likely to work	del show eximate del show eximate del show eximate where other eximate eximate eximate eximate eximple and eximply and eximply and eximply and eximply and eximate eximate eximply and eximate eximate eximply and eximate exi	tremely small variable transport of the state of the case of the c	dues for most featured lay don't matter must be a so regression also any penalty, so it is maken data. Last gnores unimportant in the overfitting bedie at the source of a source of	ch. On the other han of thinks the same that thinks the same that the sa	d, for lasso regression as alary is important in if some features  vever, this can be a gher Mean Squared form better on new eatures, even the ight not fit the training the highest error rate erage MSE is in the del while other foldowing the folds. Overation of the composition o
2.6. Analyze the Company of the Conflicted of th	lifferences be near regression molary is very importar to coefficient when the coefficient when the features.  In linear regression only for linear regression erfitting, meaning it the training data a ing.  In offers highly accurate the composition of the training data ing.  In offers highly accurate the composition of the composition	del show eximate del show eximate del show eximate where other eximate eximate eximate eximate eximple and eximply and eximply and eximply and eximply and eximate eximate eximply and eximate eximate eximply and eximate exi	tremely small value from the statures without restance are zero. Lading and the features without restant features by zero, which mea work well on new But because it is aining, but it carel by removing for new data because it is aining, but it carel by removing for new data because it is aining, but it carel by removing for new data because it is aining the state at a control of t	allues for most feature and the state of the	ch. On the other han of thinks the same that thinks the same that the sa	d, for lasso regression assalary is important assalary is important in if some features.  Wever, this can be a gher Mean Squared form better on new eatures, even the right not fit the training and the fit of the deliverage of the deliverage of the deliverage of the folds. Over a state of the deliverage of the folds. Over a state of the deliverage of the folds. Over a state of the deliverage of the folds. Over a state of the deliverage of the folds. Over a state of the deliverage of the folds. Over a state of the folds. Over a st
2.6. Analyze the Coefficients in the immension of the imm	liferences be mear regression mo lary is very importance of the control of the co	del show eximate where other features and where other features and where other features are other features and where other features are other features are other features and moderate and moderate are other forms of the fold and fold are other features are other fold are other features are other fold are o	tremely small variance or some deviation in the state of the condicates instabiling that it stands (respectively better perform from a light of the state of the condicates instabiling that it stands (respectively better perform hown fairly mixed in the state of the condicates instabiling that it stands (respectively better perform hown fairly mixed in the state of the condicates instabiling that it stands (respectively better perform hown fairly mixed in the stands (respectively better perform hown fairly mixed in t	dex]  **MAKE MODIFICAT  **MAKE MODIFICAT  **StockOptionLev  **MAKE MODIFICAT  **MAKE MODIFICAT  **StockOptionLev  **MAKE MODIFICAT  **StockOptionLev  **MAKE MODIFICAT  **StockOptionLev  **MAKE MODIFICAT  **StockOptionLev  **MAKE MODIFICAT  **MAKE MODIFICAT  **StockOptionLev  **MAKE MODIFICAT  **MAKE MODIFICAT  **MAKE MODIFICAT  **StockOptionLev  **MAKE MODIFICAT  **MAKE	ch. On the other han on thinks the same that	d, for lasso regressist salary is important in its salary is important in i

poly4\_model = make\_pipeline(PolynomialFeatures(4), LinearRegression())

poly3\_model = make\_pipeline(PolynomialFeatures(3), LinearRegression())

print(f"Mean Squared Error for 10th-degree Polynomial: {mse\_poly10:.2f}")
print(f"Mean Squared Error for 4th-degree Polynomial: {mse\_poly4:.2f}")
print(f"Mean Squared Error for 3rd-degree Polynomial: {mse\_poly3:.2f}")

print(f"\nNormalized Mean MSE across all models: {normalized\_mean\_mse:.4f}")

print(f"Normalized MSE scores for each model: {normalized\_mse\_scores}")

Mean Squared Error for 10th-degree Polynomial: 13751906782570242.00

Mean Squared Error for 4th-degree Polynomial: 8241278442.32 Mean Squared Error for 3rd-degree Polynomial: 8185296431.11

print(f"Normalized Variance of MSE across all models: {normalized\_variance\_mse:.4f}")

Normalized MSE scores for each model: [1.0, 5.992825993242384e-07, 5.952117448529439e-07]

The performance of polynomial regression models shows that the 10th degree model is better than the 4th degree and 3rd degree in terms of mean squared error (MSE). This is because the 10th degree polynomial was overfitted to the data and that automatically guided it towards poor performance on the test based set. The reduced MSE and similar performance of the 4th degree and 3rd degree models indicate a better generalization with the use of fewer parameters, as well as stability regarding hyper-parameter tuning. In normalized form, the 10th degree model was taken as a baseline (1.0) and all of the other models had an MSE that normalized to zero indicating their far greater performance compared to the 10th degree model. The mean of 0.3333 shows the normalized MSE across the models, and we see from the variance of 0.2222 that there is a great deal of difference between model performances In general, less complex polynomial models (3rd and 4th degrees) were found to be more accurate and less sensitive

to random variations in data than a highly complex model 10th degree, even though the complexities were equipped through an

 ${\it \# Normalize MSE scores by dividing by the maximum MSE (makes it easier to interpret)}$ 

poly4\_model.fit(X\_train, y\_train)

poly3\_model.fit(X\_train, y\_train)

average\_mse = np.mean(mse\_scores)
variance\_mse = np.var(mse\_scores)

max\_mse = max(mse\_scores)

# Display normalized results

y\_pred\_poly4 = poly4\_model.predict(X\_test)

# 6.5. 3th-degree polynomial regression

y\_pred\_poly3 = poly3\_model.predict(X\_test)

mse\_poly4 = mean\_squared\_error(y\_test, y\_pred\_poly4)

mse\_poly3 = mean\_squared\_error(y\_test, y\_pred\_poly3)

# 6.6. Display MSE for different polynomial degrees

mse\_scores = [mse\_poly10, mse\_poly4, mse\_poly3]

normalized mean mse = average mse / max mse

print(f"Normalized MSE = {mse:.4f}")

Cross-validation results (Normalized):

automatic search process applied beforehand.

Normalized Mean MSE across all models: 0.3333

Normalized Variance of MSE across all models: 0.2222

Normalized MSE = 1.0000 Normalized MSE = 0.0000 Normalized MSE = 0.0000

6.6. Analysis

normalized\_mse\_scores = [mse / max\_mse for mse in mse\_scores]

normalized\_variance\_mse = variance\_mse / (max\_mse \*\* 2)

print("\nCross-validation results (Normalized):")
for i, mse in enumerate(normalized\_mse\_scores, 1):

