EXPLANATION OF AWS LAMBDA RECEIPT PROCESSING CODE

This document explains the AWS Lambda function code that automates receipt processing. It uses AWS services to extract text from receipt images, store it in DynamoDB, and send email notifications.

1. AWS Services Used

The function utilizes the following AWS services:

- S3 for storing receipt images.
- Textract for OCR-based text extraction.
- DynamoDB for storing extracted data.
- SES for sending email notifications.
- IAM for secure access management.

2. lambda_handler Function

The `lambda_handler` function is the entry point for the Lambda. It handles the following steps:

```
def lambda_handler(event, context):
    # 1. Retrieve S3 bucket and key
    bucket = event['Records'][0]['s3']['bucket']['name']
    key = urllib.parse.unquote_plus(event['Records'][0]['s3']['object']['key'])
```

It then verifies the object, processes it, stores the data, and sends an email.

3. process_receipt_with_textract Function

This function calls Textract to analyze the receipt and extract data:

```
response = textract.analyze_expense(
    Document={ 'S3Object': { 'Bucket': bucket, 'Name': key } }
)
```

It then builds a dictionary with `receipt_id`, `vendor`, `total`, `date`, and `items`.

4. store_receipt_in_dynamodb Function

This function writes the extracted data into DynamoDB:

```
table = dynamodb.Table(DYNAMODB_TABLE)
db_item = {
    'receipt_id': receipt_data['receipt_id'],
    'vendor': receipt_data['vendor'],
    'total': receipt_data['total'],
    'items': items_for_db,
    'processed_timestamp': datetime.now().isoformat()
}
table.put_item(Item=db_item)
```

5. send_email_notification Function

This function sends an email using SES with the receipt summary:

```
ses.send_email(
    Source=SES_SENDER_EMAIL,
    Destination={ 'ToAddresses': [SES_RECIPIENT_EMAIL] },
    Message={ 'Subject': {'Data': f"Receipt Processed:
{receipt_data['vendor']}"},
            'Body': {'Html': {'Data': html_body}} }
)
```

6. Conclusion

This AWS Lambda function implements a fully automated receipt processing workflow using AWS services. The code extracts text from receipt images, stores it in DynamoDB, and notifies users via email, eliminating manual data entry.